

Release 2.1D John F. Collins, BioComputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Tue Feb 3 19:21:38 1998; MasPar time 568.41 Seconds
Abular output not generated. 1157.009 Million cell updates/sec

Title: >US-08-713-928A-5
Description: (1-463) from US08713928A.seq
Perfect Score: 463
N.A. Sequence: 1 CAATACGATATTACCAATA.....CGGTGAAGTAGCAGSNC 463
Comp: GTTAGCTATATGGCTTAT.....GGCCACTTCATTCGTCWAG

Scoring table:
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 430261 seqs, 710217276 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb1-new7

9: HUM2 10: HUM3 11: INV1 12: INV2 13: ORG 14: MAM 15: VRT
16: PLN 17: PRO1 18: PRO2 19: ROD 20: SYN 21: UNC 22: VIR
genbankal01

23: BCT1 24: BCT2 25: BCT3 26: BCT4 27: BCT5 28: BCT6 29: BCT7
30: BCT8 31: BCT9 32: BCT10 33: BCT11 34: BCT12 35: BCT13
36: GEN1 37: GEN2 38: GEN3 39: GEN4 40: GEN5 41: GEN6 42: HTG1
43: HTG2 44: HTG3 45: HTG4 46: HTG5 47: INV1 48: INV2 49: INV3
50: INV4 51: INV5 52: INV6 53: INV7 54: INV8 55: INV9 56: INV10
57: INV11 58: INV12 59: MAM1 60: MAM2 61: MAM3 62: VRT1
63: VRT2 64: VRT3 65: VRT4 66: PAT1 67: PAT2 68: PAT3 69: PAT4
70: PAT5 71: PAT6 72: PAT7 73: PHG 74: PLN1 75: PLN2 76: PLN3
77: PLN4 78: PLN5 79: PLN6 80: PLN7 81: PLN8 82: PLN9 83: PLN10
84: PLN11 85: PLN12 86: PLN1 87: PRI2 88: PRI3 89: PRI4
90: PRI5 91: PRI6 92: PRI7 93: PRI8 94: PRI9 95: PRI10
96: PRI11 97: PRI12 98: PRI13 99: PRI14 100: PRI15 101: PRI16
102: PRI17 103: ROD1 104: ROD2 105: ROD3 106: ROD4 107: ROD5
108: ROD6 109: ROD7 110: ROD8 111: ROD9 112: STR 113: SYN
114: UNA
genbankb101
115: VRL1 116: VRL2 117: VRL3 118: VRL4 119: VRL5 120: VRL6
121: VRL7 122: VRL8 123: VRL9 124: VRL10 125: VRL11
genbank-new7
126: BCT 127: GEN 128: HTG1 129: HTG2 130: INV 131: MAM
132: VRT 133: PHG 134: PLN1 135: PLN2 136: PRI1 137: PRI2
138: ROD 139: SYN 140: UNA 141: VRL
u-emb151_101
142: part1 143: part2

Statistics: Mean 10.942; Variance 9.088; scale 1.204

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	428	92.4	16720	134	LEU68072	Lycopersicon esculent	1.91e-194
2	428	92.4	16720	134	LEU68072	Lycopersicon esculent	1.91e-194
3	294	63.5	2884	84	TOMHMG2A	Tomato 3-hydroxy-3-me	2.35e-126
4	156	33.7	184	78	LEU68072	Lycopersicon esculent	1.58e-57
5	36	7.8	1095	76	DDICSA	Dictyostelium discoid	2.02e-02
6	36	7.8	3700	51	DDGP80G	D.discoideum gp80 gen	2.02e-02
7	33	7.1	14001	57	PFCOMP1RB	P.falciiparum complete	2.68e-01
8	32	6.9	2379	51	DD31	D.discoideum culminat	6.22e-01
9	31	6.7	1124	39	BDU2074	F25H9-T7.2 IGF Arabid	1.43e+00
10	31	6.7	1891	51	DDU73685	Dictyostelium discoid	1.43e+00
11	31	6.7	6615	78	MISDCYTB	S.douglasii gene for	1.43e+00
12	31	6.7	36238	43	CEK01G12	*** SEQUENCING IN PRO	1.43e+00
13	31	6.7	50751	42	CEFO2H6	*** SEQUENCING IN PRO	1.43e+00
14	31	6.7	11062	46	HSAC001237	*** SEQUENCING IN PRO	1.43e+00
15	30	6.5	192	16	LEU68071	Lycopersicon esculent	3.24e+00
16	30	6.5	192	78	LEU68071	Lycopersicon esculent	3.24e+00
17	30	6.5	192	134	LEU68071	Lycopersicon esculent	3.24e+00
18	30	6.5	854	78	MITGTREN6	Torulopsis glabrata m	3.24e+00
19	30	6.5	854	85	YSLMTIG08	Yeast (T.glabrata) ml	3.24e+00
20	30	6.5	865	57	PLU39369	Parastenopa limata 16	3.24e+00
21	30	6.5	1868	125	VSVGLYPO	Vesicular stomatitis	3.24e+00
22	30	6.5	33010	50	CELZC196	Caenorhabditis elegans	3.24e+00
23	29	6.3	116	66	A08900	H.sapiens (haplotype	7.26e+00
24	29	6.3	876	39	B12885	T17A8-T7 TAMU Arabido	7.26e+00
25	29	6.3	3000	77	DDIGF24	D.discoideum glycopro	7.26e+00
26	29	6.3	3075	58	XPFNAPOL	P.falciiparum gene for	7.26e+00
27	29	6.3	7038	57	PFGP195A	Plasmodium falciparum	7.26e+00
28	29	6.3	8633	57	PFPMDR1	P.falciiparum pfmdr1 g	7.26e+00
29	29	6.3	15421	57	PFCOMP1RA	P.falciiparum complete	7.26e+00
30	29	6.3	22243	57	PFAVAR23A	Plasmodium falciparum	7.26e+00
31	29	6.3	28078	48	CE2F1F12	Caenorhabditis elegans	7.26e+00
32	29	6.3	37393	94	HSU232G2	Human DNA sequence fr	7.26e+00
33	29	6.3	95936	44	HS127L4	Human DNA sequence **	7.26e+00
34	29	6.3	172533	8	HSAC2381	Human BAC clone RG020	7.26e+00
35	29	6.3	200434	45	HS415G2	Human DNA sequence **	7.26e+00
36	29	6.3	201239	45	HS435D1	Human DNA sequence **	7.26e+00
37	28	6.0	512	39	B11004	F19A23-T7 IGF Arabido	1.61e+01
38	28	6.0	1147	39	B13042	T3OM24-Sp6.1 TAMU Ara	1.61e+01
39	28	6.0	1452	95	HSU63108	Human eukaryotic init	1.61e+01
40	28	6.0	1950	51	DDECA	D.discoideum ecma pro	1.61e+01
41	28	6.0	2019	84	STU51985	Solanum tuberosum clo	1.61e+01
42	28	6.0	17084	60	DWMTGNME	D.virginiana complete	1.61e+01
43	28	6.0	28951	92	HSL161C2	Human DNA sequence fr	1.61e+01
44	28	6.0	34860	50	CELR173	Caenorhabditis elegans	1.61e+01
45	28	6.0	90756	129	AC002342	*** SEQUENCING IN PRO	1.61e+01

ALIGNMENTS

RESULT 1
ID LEU68072 standard; DNA; PLN; 16720 BP.
AC U68072;
NI G2246449
DT 13-SEP-1996 (Rel. 49, Created)
DE 10-JUL-1997 (Rel. 52, Last updated, Version 2)
DE Lycopersicon esculentum copia-like retrotransposon T0RT1L,
DE 3-hydroxy-3-methylglutaryl CoA reductase 2 (HMG2) gene, complete
DE cds.
KW Lycopersicon esculentum (tomato)
OS Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
OC Charophyta/Embryophyta group; Embryophyta; vascular plants;
OC seed plants; Magnoliophyta; Magnoliopsida; Solanaceae;
OC Solanaceae; Solanum clade; Lycopersicon.
RN [1]
RP 1-16720

NID	92246449	
KEYWORDS	tomato.	
SOURCE	Lycopersicon esculentum	
ORGANISM	Eukaryotes; mitochondrial eukaryotes; Viridiplantae; Charophyta/Embryophyta group; Embryophyta; vascular plants; seed plants; Magnoliophyta; Magnoliopsida; Solananae; Solanales; Solanaceae; Solanum clade; Lycopersicon.	
REFERENCE	1 (bases 1 to 16720)	
AUTHORS	Daraseelia,N.D., Tarchevskaya,S. and Narita,J.O.	
TITLE	The promoter for tomato 3-hydroxy-3-methylglutaryl coenzyme A reductase gene 2 has unusual regulatory elements that direct high-level expression	
JOURNAL	Plant Physiol.	
REFERENCE	112 (2), 727-733 (1996)	
AUTHORS	Narita,J.O.	
TITLE	Direct Submission	
JOURNAL	Submitted (27-AUG-1996) Department of Biological Sciences, University of Illinois at Chicago, 900 S. Ashland Ave, Chicago, IL 60607, USA	
REFERENCE	3 (bases 1 to 16720)	
AUTHORS	Narita,J.O.	
TITLE	Direct Submission	
JOURNAL	Submitted (19-JUN-1997) Department of Biological Sciences, University of Illinois at Chicago, 900 S. Ashland Ave, Chicago, IL 60607, USA	
REMARK	Sequence update by submitter	
FEATURES	Location/Qualifiers	
source	1..16720	
repeat_region	/organism="Lycopersicon esculentum"	
LTR	/strain="VFNT Cherry"	
source	complement(1526..1531)	
misc_feature	/note="putative insertion site duplication"	
CDS	complement(1532..2341)	
	/note="3' LTR"	
	complement(1532..11220)	
	/organism="Lycopersicon esculentum"	
	/note="copla-like retrotransposon ToRTLI"	
	complement(2342..2355)	
	/note="polypurine tract"	
	complement(2769..4592)	
	/note="ORF 3"	
	/codon_start=1	
	/pseudo	
	complement(5092..7545)	
	/note="ORF 2"	
	/codon_start=1	
	/pseudo	
	complement(7542..10349)	
	/note="ORF 1; putative"	
	/codon_start=1	
	/pseudo	
primer_bind	complement(10397..10408)	
LTR	/note="putative primer binding site"	
	10412..11220	
repeat_region	/note="5' LTR"	
TATA_signal	complement(11221..11226)	
5'UTR	/note="putative insertion site duplication"	
	12627..12641	
gene	12664..12771	
	/gene="HMG2"	
	12664..15576	
	/gene="HMG2"	
exon	12664..13805	
	/gene="HMG2"	
mRNA	join(12664..13805,13888..14069,14272..14618,15034..>15576)	
	/gene="HMG2"	
CDS	/product="3-hydroxy-3-methylglutaryl CoA reductase 2"	
	join(12772..13805,13888..14069,14272..14618,15034..15276)	
	/gene="HMG2"	
	/codon_start=1	

	Dbb	181	ctacatccttctctcccatcccatacactctcttttaaacaattatatctgtccaat	240
	QY	325	CATACATTCTTCTCCTCACATCCCATCACTCTCTTTTAACAATTATACTTGTCAT	384
	Ddb	241	cataatcccccaaaaacaaacctttttctctctctctcttcaccgcggcgagacttac	300
	QY	385	CATCAATCCCACAAACAACACTTTTTTCTCTCTCTCTCTACCAGCGGCGAGACTTAC	444
	Ddb	301	cggtga aa 308	
	QY	445	CGGTGA AA 452	
RESULT	4			
LOCUS	LEU68072	184 bp	DNA	PLN 12-SEP-1996
DEFINITION	Lycopersicon esculentum 3-hydroxy-3-methylglutaryl CoA reductase	(HMG2) gene, partial cds and 5' untranslated region.		
ACCESSION	U68072			
NID	g1532207			
KEYWORDS	.	tomato.		
SOURCE	Lycopersicon esculentum			
ORGANISM	Eukaryotes; mitochondrial eukaryotes; Viridiplantae; Charophyta/Embryophyta group; Embryophyta; Magnoliophyta; Magnoliopsida; Solananae; Solanales; Solanaceae; Solanum clad-			
REFERENCE	1 (bases 1 to 184)			
AUTHORS	Dasaralia,N.D., Tarchevskaya,S. and Narita,J.O. The Promoter for Tomato 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase Gene 2 Has Unusual Regulatory Elements That Direct High-level Expression Plant Physiol. 112 (1996) In press			
JOURNAL	2 (bases 1 to 184)			
REFERENCE	Narita,J.O. Direct Submission Submitted (27-AUG-1996) Department of Biological Sciences, University of Illinois at Chicago, 900 S. Ashland Ave, Chicago 60607, USA			
AUTHORS				
TITLE				
JOURNAL				
FEATURES	Location/Qualifiers			
SOURCE	1..184	/organism="Lycopersicon esculentum"		
TATA_signal	28...42			
mRNA	63...184	/gene="HMG2" /product="3-hydroxy-3-methylglutaryl CoA reductase 2"		
gene	63...184	/gene="HMG2" /note="HMGR2"		
CDS	173...>184	/codon_start=1 /product="3-hydroxy-3-methylglutaryl CoA reductase 2"		
BASE COUNT	47 a 59 c 21 g 57 t			
ORIGIN				
Query Match	33.7%; Score 156;	DB 78; Length 184;		
Best Local Similarity	98.8%;	Pred.No.1.58e-57;		
Matches	168; Conservative	0; Mismatches 0; Indels 2; Gaps		
Ddb	1	aagtccaggcgga accgggttcctcataaatcacatttcctacattctctctcc	60	
QY	285	AAAGTCCAGCGGGCAACGGGTTCT--ATAAATACATTCTCTACATCTTCTCTCTCC	342	
Ddb	61	tcaatccccatcacctcttctttaacaattatacttgtaacatcatsaacatccacaaaa ca	120	
QY	343	TCACATCCCCATCACTCTCTTTTAACAATTAATCTTGTCATCATCAATCCCAAAA CA	402	
Ddb	121	caactttttctctctcttttttcta ccaggcgga acttac cggtga aa 170		
QY	403	CAC TT TTT CT C TC T C T T T T T T C T C A C C G G G C G A C Y T A C C G T G A A	452	

[illegible]

Best Available Copy

[illegible]

~~SECRET~~
Note: remainder of annotations omitted.

LOCUS	BI2074	1124 bp	DNA	GSS	13-MAY-1997
DEFINITION	F53H9-T7.2 IGF Arabidopsis thaliana genomic clone F53H9.				
ACCESSION	BI2074				
VERSION	g2093194				
NID	g2093194				
KEYWORDS	GSS.				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
	Eukaryotae; Mitochondrial eukaryotes; Viridiplantae;				

[illegible]

```

QY 219 AAATTAAGAAAAAGAAAGTATATATTGTAAAAAGATAACTCCATCAAAATATAAA 278
      887
      || || |||
QY 279 ATGAAAAAA 287

RESULT 15
ID LEU68071 standard; DNA; PLN; 192 BP.
AC U68071;
NI g1532206
DT 13-SEP-1996 (Rel. 49, Created)
DT 10-JUL-1997 (Rel. 52, Last updated, Version 4)
DE Lycopersicon esculentum 3-hydroxy-3-methylglutaryl-CoA reductase 1
DE (HMG1) gene, partial cds and 5' untranslated region.
KW (HMG1)
OC Lycopersicon esculentum (tomato)
OC Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
OC Charophyta/Embryophyta group; Embryophyta; vascular plants;
OC seed plants; Magnoliophyta; Magnoliopsida; Solanaceae; Solanaceae;
OC Solanaceae; Solanum clade; Lycopersicon.
RN [1]
RP 1-192
RA Daraselia N.D., Tarchevskaya S., Narita J.O.;
RP "The promoter for tomato 3-hydroxy-3-methylglutaryl coenzyme A
RT reductase gene 2 has unusual regulatory elements that direct
RT high-level expression";
RT Plant Physiol. 112:727-733(1996).
RL [2]
RP 1-192
RA Narita J.O.;
RT ;
RRL Submitted (27-AUG-1996) to the EMBL/GenBank/DBJ databases.
RRL Department of Biological Sciences, University of Illinois at
RRL Chicago, 900 S. Ashland Ave, Chicago, IL 60607, USA
RFH Key
RFH Location/Qualifiers
FH source
FH 1..192
FT /organism="Lycopersicon esculentum"
FT TATA_signal
FT 37..51
FT mRNA
FT 72..192
FT /gene="HMG1"
FT /product="3-hydroxy-3-methylglutaryl-CoA reductase 1"
FT 181..>192
FT /gene="HMG1"
FT /note="HMGR1"
FT /codon_start=1
FT /product="3-hydroxy-3-methylglutaryl-CoA reductase 1"
FT /db_xref="PID:e265716"
FT /translation="MDVR"
FT Sequence 192 BP; 46 A; 54 C; 20 G; 72 T; 0 other;
SQ

Query Match 6.5%; Score 30; DB 16; Length 192;
Best Local Similarity 81.3%; Pred. No. 3.24e+00;
Matches 39; Conservative 0; Mismatches 9; Indels 0; Gaps

Db 37 ctataaatacattttctactttctgcttccacacaaacatcact 84
|||||
QY 310 CTATAAATACATTTCTTACATCTTCTCTCTCCCTCACATCCCATCACT 357

Search completed: Tue Feb 3 19:44:28 1998
Job time : 1370 secs.

```

MPsrc_h_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligenetics, Inc.

MPsrc_h_nn n.a. - n.a. database search, using Smith-Waterman algorithm

on on: Tue Feb 3 19:44:49 1998; MasPar time 66.98 Seconds
Fabular output not generated. 797.669 Million cell updates/sec

Title: >US-08-713-928A-5
Description: (1-463) from US08713928A.seq
Perfect Score: 463
N.A. Sequence: 1 CAATACGATATACCGAATA.....CCGCTGAAGTAAGCAGSTC 463
Comp: GTATGCTATATAGCTTAT.....GCCCACTTCATTCGTCWAG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 159651 seqs, 57698962 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq30
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33

Statistics: Mean 8.315; Variance 7.573; scale 1.098

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description	Pred. No.
1	463	100.0	463	32	7T1752	1.04e-198
2	434	93.7	1388	13	Q81473	4.13e-185
3	45	9.7	1047	2	Q10572	1.47e-07
4	42	9.1	1047	2	Q10572	2.29e-06
5	41	8.9	204	1	N81164	5.66e-06
6	35	7.6	91	9	Q51746	1.18e-03
7	35	7.6	204	1	N81164	1.18e-03
8	32	6.9	91	9	Q51746	1.00e-00
9	27	5.8	68	32	T63255	1.57e-02
10	27	5.8	68	32	T63255	1.00e-00
11	27	5.8	70	24	T14325	1.00e-00
12	27	5.8	7400	17	T09340	1.00e-00
13	27	5.8	9789	24	T41852	1.00e-00
14	26	5.6	114	12	Q70467	2.25e+00
15	26	5.6	140	31	T76781	2.25e+00

16	26	5.6	240	31	T76782	Staphylococcus aureus	2.25e+00
17	25	5.4	67	24	T14322	Primer used in the la	4.96e+00
18	25	5.4	68	32	T63255	Messenger RNA primer	4.96e+00
19	25	5.4	68	33	T73397	Oligonucleotide tag c	4.96e+00
20	25	5.4	70	24	T14325	Conjugate formed by l	4.96e+00
21	25	5.4	114	12	Q70472	Generic DNA sequence	4.96e+00
22	25	5.4	114	12	Q70465	Generic DNA sequence	4.96e+00
23	25	5.4	114	12	Q70469	Generic DNA sequence	4.96e+00
24	25	5.4	688	18	T16856	Integrin subunit beta	4.96e+00
25	25	5.4	1046	2	Q10377	Plasmid pMG3C9 used t	4.96e+00
26	25	5.4	1173	1	Q06504	CbEPV spheroidin gene	4.96e+00
27	25	5.4	1470	10	Q55185	MS-Lel610 Vector.	4.96e+00
28	25	5.4	1582	22	T28259	Survival motor neuron	4.96e+00
29	25	5.4	1582	21	T18831	Human survival motor	4.96e+00
30	25	5.4	1754	3	Q20239	Gene encoding cell gr	4.96e+00
31	25	5.4	2504	31	T73286	S. pombe origin of re	4.96e+00
32	25	5.4	2504	30	T62359	Schizosaccharomyces p	4.96e+00
33	25	5.4	2960	29	T41616	Adzuki bean endo-xylo	4.96e+00
34	25	5.4	5852	2	Q11710	Dictyostelium plasmid	4.96e+00
35	25	5.4	19124	32	T72882	Plasmodium var-7 gene	4.96e+00
36	24	5.2	114	12	Q70465	Generic DNA sequence	1.08e+01
37	24	5.2	903	29	T47670	Human G protein gamma	1.08e+01
38	24	5.2	1594	27	T35787	Human retinoid X rece	1.08e+01
39	24	5.2	1878	1	N90416	DNA sequence encoding	1.08e+01
40	24	5.2	2245	20	T11023	Urochloa panicoides C	1.08e+01
41	24	5.2	3101	1	Q02047	Sequence encoding a c	1.08e+01
42	24	5.2	3905	4	Q25893	Human Factor XIII cod	1.08e+01
43	24	5.2	6020	1	Q06648	Plasminogen gene from	1.08e+01
44	24	5.2	53577	28	T18551	Human polycystic kidn	1.08e+01
45	24	5.2	580073	27	T58840	Mycoplasma genitalium	1.08e+01

ALIGNMENTS

RESULT 1
ID T71752 standard; DNA; 463 BP.
AC T71752;
DT 01-OCT-1997 (first entry)
DE Mega promoter.
KW Lysosomal enzyme; lysosome; transgenic plant; glucocerebrosidase;
KW alpha-L-iduronidase; IUDA; enzyme replacement therapy;
KW Gaucher disease; Hurler syndrome; Mega promoter; ss.
OS Synthetic.
PN WO9710353-A1.
PD 20-MAR-1997.
PF 13-SEP-1996; U14730.
PR 14-SEP-1995; US-003737.
PA (CROP-) CROPTech DEV CORP.
PA (VIRG) VIRGINIA TECH INTELLECTUAL PTY INC.
PI Cramer CL, Oishi KK, Radin DN, Weissenborn DL;
DR WPI; 97-202248/18.
PT Production of enzymatically active (modified) lysosomal enzyme in
PT transgenic plants - useful in treatment of lysosomal storage
PT disorders
PS Claim 4; Page 64; 111pp; English.
CC The Mega promoter (T71752), a modified tomato HMG2 promoter, has a
CC low basal expression in unstressed plant tissues, but is highly
CC induced in both immature and mature tissues by the process of
CC mechanical gene activation, or by chemicals that induce plant
CC defence responses. Novel plant-based expression systems utilise
CC the Mega promoter to control expression of enzymatically active
CC (modified) lysosomal enzymes in transgenic plants. Such enzymes
CC include human glucocerebrosidase (see also T71753) and human alpha-
CC L-iduronidase (see also T71754). Plant expression provides for
CC post-translational modification and processing to provide enzymes
CC suitable for use in enzyme replacement therapy of lysosomal storage
CC diseases.
SQ Sequence 463 BP; 172 A; 96 C; 38 G; 156 T;

Query Match 100.0%; Score 463; DB 32; Length 463;
Best Local Similarity 100.0%; Pred. No. 1.04e-198;
Matches 463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 caatcagatattaccgaatattatactaaatacaaaatttaatttatacatcaattat 60
QY 1 CAATACGATATTACCGAATATTATATACTAAATCAAAATTTAATTATCATCATCAATTATTA 60
Db 61 aactgatatttcaaaatttcaatttatactatacttcaacttattacacaaattatc 120
QY 61 AACTGATATTCAAAATTTTAAATATTAAATATCTACATTTCAACTATTATTACCTAATTATC 120
Db 121 aaatcaaaatgatagattatttcaatatactagcccggttcgtatccaaatattttacact 180
QY 121 AAATCAAAATGATAGATTATTTCATAATAGCCCGAGTTCGTATCCAAATATTATTACACT 180
Db 181 tcacagtcacacttactataaaacttacttcaaaataataaaataaaagaaagta 240
QY 181 TCACAGTCACACTTGACTATATAAACTTTACTTCAAAATTTAAAAAAGAAAGTA 240
Db 241 tattattgaaaagataaactcctcctcaaaataataaaatgaaaagtcagcgcgca 300
QY 241 TATTATTGTAAGATAAATACCTCCATTCATTAATATAAATGAAGTCCAGCGGGCA 300
Db 301 accgggttcctataaatacttctcctacatcttctctcctcctcctcctcctcct 360
QY 301 ACCGGGTTCCTATAATACATTTCTTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Db 361 cttttaacaattatacttgcataatcaatcccaacaaacaaacaaacttttctctctct 420
QY 361 CTTTAAACAATTATATTCTGCAATCATCAATCCCAACAAACAACTTTTCTCTCTCTCT 420
Db 421 ttctctcaccgcgccagacttaccggtgaaagtaagcagstc 463
QY 421 TTTCTCACCAGCGGCGAGACTTACCAGGTGAAGTAAGCAGSTC 463

RESULT 2
ID Q81473 standard; DNA; 1388 BP.
AC Q81473;
DT 04-SEP-1995 (first entry)
DE HM2 promoter 1.
KW HM2; promoter; transgenic plant; post-harvest production;
KW 3-hydroxy-3-methylglutaryl-CoA-reductase; HMGR; wounding;
KW pathogen infection; pest infestation; inducible expression;
KW pollen-specific expression; disease-resistance;
KW pesticide-resistance; crop improvement; pdw101; ss.
KW *Lycopodium obscurum* cv. VFNT Cherry.
FH Key Location/Qualifiers
FT misc RNA 1037
FT /tag= a
FT /note= "transcriptional start site"
FT TATA_signal 1003..1010
FT /tag= b
FT WO9503690-A.
PD 09-FEB-1995.
PE 02-AUG-1994; U08722.
PR 02-AUG-1993; US-100816.
PA (VIRG) VIRGINIA TECH INTELLECTUAL PROPERTIES.
PI Cramer CL, Weissenborn DL;
DR WPI; 95-081942/11.
PT HM2 promoter expression systems - used for inducible prodn. of
PT gene prods. in plants and plant cell cultures, partic.
PT post-harvest prodn.
PS Disclosure; Fig. 4a-4c; 110pp; English.
CC A 2.5 kb EcoRI fragment of the HMGR gene (HM2)-containing clone
CC pTH295 was inserted into the EcoRI site of Bluecript SK-vector to
CC generate pdw101. A 1388 bp sequence from the 3' end of the 2.5 kb
CC EcoRI insert of pdw101 is given in Q81473. This promoter region
CC can be used to control gene expression in transgenic plants or
CC cell cultures in response to wounding, elicitors, pest infestation,
CC pathogen infection, etc.
SQ Sequence 1388 BP; 490 A; 238 C; 168 G; 492 T;
Query Match 93.7%; Score 434; DB 13; Length 1388;
Best Local Similarity 99.3%; Pred. No. 4.13e-185;
Matches 452; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

Db 590 caatcagatattaccgaatattatactaaatacaaaatttatacatcaattat 749
QY 1 CAATACGATATTACCGAATATTATATACTAAATCAAAATTTAATTATCATCAATTATTA 59
Db 750 aaactgatatttcaaaatttcaatttatactatacttcaacttattacacaaattat 809
QY 60 AAATCAAAATGATAGATTATTTCATAATAGCCCGAGTTCGTATCCAAATATTATTAC 119
Db 810 caaatcaaaatgatagattatttcaatatactagcccggttcgtatccaaatattttacac 869
QY 120 CAAATCAAAATGATAGATTATTTCATAATAGCCCGAGTTCGTATCCAAATATTATTAC 179
Db 870 ttgacagtcacacttactataaaacttacttcaaaataataaaataaaagaaagta 929
QY 180 TTGACAGTCACACTTGACTATATAAACTTTACTTCAAAATTTAAAAAAGAAAGT 239
Db 930 atattattgaaaagataaactcctcctcaaaataataaaatgaaaagtcagcgcgca 989
QY 240 ATATTATTGTAAGATAAATACCTCCATTCATTAATATAAATGAAGTCCAGCGGGC 299
Db 990 aacgggttcctataaatacttctcctacatcttctctcctcctcctcctcctcct 1049
QY 300 AACCGGGTTCCT--ATAATACATTTCTTACATCTCTCTCTCTCTCTCTCTCTCTCTCT 357
Db 1050 cttcttcaacaattatacttgcataatcaatcccaacaaacaaacttttctctct 1109
QY 358 CTTCTTTTAAACAATTATATTCTGCAATCATCAATCCCAACAAACAACTTTTCTCTCTCT 417
Db 1110 ctttctcaccgcgccagacttaccggtgaaagtaagcagstc 1144
QY 418 CTTTCTCACCAGCGGCGAGACTTACCAGGTGAAGTAAGCAGSTC 452

RESULT 3
ID Q10572 standard; DNA; 1047 BP.
AC Q10572;
DT 09-APR-1991 (first entry)
DE Human Natriuretic Peptide Receptor B.
KW NPNB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW hyperaldosteronism; glaucoma; guanyl cyclase.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..22
FT /label= signal sequence
FT Protein 12
FT /label= mature NPBR
FT Domain 23..455
FT /label= extracellular domain
FT /note= "binds natriuretic peptides A,B and C1"
FT Domain 456..456
FT /label= transmembrane domain
FT Domain 479..1047
FT /label= cytoplasmic domain
FT /note= "GC and protein kinase activity"
FT Modified-site 24..26
FT /label= N-glycosite
FT Modified-site 35..37
FT /label= N-glycosite
FT Modified-site 161..163
FT /label= N-glycosite
FT Modified-site 195..197
FT /label= N-glycosite
FT Modified-site 244..246
FT /label= N-glycosite
FT Modified-site 277..279
FT /label= N-glycosite
FT Modified-site 349..351
FT /label= N-glycosite
FT Modified-site 600..602
FT /label= N-glycosite
PN WO9100292-A.
PD 10-JAN-1991.

PR 23-JUN-1990; U03586.
 PR 23-JUN-1989; US-370673.
 PA (GETH) GENENTECH INC.
 PI Chang M, Goeddel D, Lowe D;
 DR WPI; 91-036711/05.
 DR N-PSDB; Q10324.
 PT Natriuretic protein receptor B - for diagnosis and treatment of
 PT kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
 PS Claim 3; Fig 1; 49pp; English.
 CC The sequence was derived from the DNA encoding natriuretic peptide
 CC receptor B, NPRB, having guanyl cyclase (GC) activity and protein
 CC kinase activity. The DNA can be inserted into expression vectors
 CC for the prodn. of the protein, opt. after being mutated to produce
 CC NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
 CC 114,952). The protein (or variants) can be used in treatment of
 CC natriuretic peptide disorders, and also to isolate peptides using
 CC affinity chromatography. Antibodies with affinity for NPRB can
 CC also be prep'd.
 SQ Sequence 1047 BP; 87 A; 15 C; 83 G; 51 T;
 Query Match 9.7%; Score 45; DB 2; Length 1047;
 Best Local Similarity 9.8%; Pred. No. 1.47e-07;
 Matches 41; Conservative 111; Mismatches 274; Indels 3; Gaps 3;
 Db 471 gyssnnnnknnnnknaasmvrvnnnnngnsnryhkgagsrntn-snrqssygsnmt 529
 Cp 427 GAGGAAAAGAGAGAGAGAAAAGTGTGTTGGGATGATGATGATGATGATGATG 368
 Db 530 abgkyannantghkgnvvankhvkrrnntrnnvnnnnkhmrdrnnnnhtrnnngacndn 589
 Cp 367 TTAAGAAGAGAGTGGGATGAGGAGAGAGAGATGAGGAAA-TCATATTATAGG 309
 Db 590 nncnvtncnrgsnnndnnnnndnnnnrnsnnndvnmrvnnnnhnsnshgsksn 649
 Cp 308 AACCCGGTTCGCGCGTGCACITTTTTCATTTTATATTTGAATGAGTATATCTTTA 249
 Db 650 cvvdrnvnknidynasnrstannndnnanyaknnntannnnsgnnntgmnaadvys 709
 Cp 248 CAATAATATACITTCITTCITTTTATTTTGAAGTAAAGTTTATATACAAAGTTG 189
 Db 710 ngnnnnnnaursgnyngndnsnknvknvngnrynnrnsndrtnnnnnvnmnrw 769
 Cp 188 ACTGTCAGTGTAAATATTTGGATACGACTGGCTATTATGAATACTCATACATT 129
 Db 770 andanrndngnknrrnnknngtssndnnnnrnmnyannnnknvnrtnaynnknrk 829
 Cp 128 TTGCATTTCATAATTAGGTAATAATAGTGAAGTAGATATTAATAATTAATAATT-GAA 70
 Db 830 anannynnnhsvannknrgntvnanandsvtynsdvngntansanstmnnvvtnnndn 889
 Cp 69 ATATCAGTTTAATAATTGATGATGATAAATAATTTTATTTGATGATGATGATGAT 10
 Db 890 ytcndannnd 898
 Cp 9 ATCGATTG 1
 RESULT 4
 ID Q10572 standard; DNA; 1047 BP.
 AC Q10572;
 DT 09-APR-1991 (first entry)
 DE Human Natriuretic Peptide Receptor B.
 KW NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
 KW hyperaldosteronism; glaucoma; guanyl cyclase.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1..22
 FT /label= signal sequence
 FT /label= protein 12
 FT /label= mature NPRB 23..455
 FT Domain
 FT /label= extracellular domain
 FT /note= "binds natriuretic peptides A,B and C]"

FT Domain 456..456
 FT /label= transmembrane domain
 FT Domain 479..1047
 FT /label= cytoplasmic domain
 FT /note= "GC and protien kinase activity"
 FT Modified-site 24..26
 FT /label= N-glycos-site
 FT Modified-site 35..37
 FT /label= N-glycos-site
 FT Modified-site 161..163
 FT /label= N-glycos-site
 FT Modified-site 195..197
 FT /label= N-glycos-site
 FT Modified-site 244..246
 FT /label= N-glycos-site
 FT Modified-site 277..279
 FT /label= N-glycos-site
 FT Modified-site 349..351
 FT /label= N-glycos-site
 FT Modified-site 600..602
 FT /label= N-glycos-site
 FT W09100292-A.
 PN 10-JAN-1991.
 PD 22-JUN-1990; U03586.
 PE 23-JUN-1989; US-370673.
 PR (GETH) GENENTECH INC.
 PA Chang M, Goeddel D, Lowe D;
 PI WPI; 91-036711/05.
 DR N-PSDB; Q10324.
 DR Natriuretic protein receptor B - for diagnosis and treatment of
 PT kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
 PS Claim 3; Fig 1; 49pp; English.
 CC The sequence was derived from the DNA encoding natriuretic peptide
 CC receptor B, NPRB, having guanyl cyclase (GC) activity and protein
 CC kinase activity. The DNA can be inserted into expression vectors
 CC for the prodn. of the protein, opt. after being mutated to produce
 CC NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
 CC 114,952). The protein (or variants) can be used in treatment of
 CC natriuretic peptide disorders, and also to isolate peptides using
 CC affinity chromatography. Antibodies with affinity for NPRB can
 CC also be prep'd.
 SQ Sequence 1047 BP; 87 A; 15 C; 83 G; 51 T;
 Query Match 9.1%; Score 42; DB 2; Length 1047;
 Best Local Similarity 8.4%; Pred. No. 2.29e-06;
 Matches 34; Conservative 109; Mismatches 258; Indels 3; Gaps 3;
 Db 547 vvankhvkrrnntrnnvnnnnkhmrdrnnnnhtrnnngacndnnnncnvtnyngnsnd 606
 QY 60 AAACATGATATTTCAAATTTTAAATATTAATATCTACTTCAACTATTATACCTAATAT 119
 Db 607 nnnndnnndwmn-rysnnnndvknmnnhnsnshgsksnkscvvdsvrnnvknndtgy 665
 QY 120 CAAATGCAAAATGATGATGATTTTCAATATAGCCCATCGTATCCCAATATTTTACAC 179
 Db 666 asnrstannndnnanyaknnntannnnsgnnntgmnaadvysngnnnnnnnrsngny 725
 QY 180 TTGACCAGTCACTTGACTATATAAACTTTACTTCAAAAATTAATAAAAGAAAGT 239
 Db 726 nngndnsnknvknvngnrynnrnsndrtnnnnnvnnnmnrwvndnndnngnknkn 785
 QY 240 ATATTATTGTAAGAGATAACTCAATCAAAATATAAATG-AAAAAAGTCCAGCGCG 298
 Db 786 nrrnknngtssndnnnnrnmnyannnnknvnrtnaynnnnknkanannynnnhsvann 845
 QY 299 CAACCGGGTTCCCTATAAATACATTTCTCATCTCTTCTCTCTCATCTCCCTACT 358
 Db 846 nkrgntvnanandsvtynsdvngntansanstmnnvvtnnndnndtnddndndvkv 905
 QY 359 TTTCTTTTAACAATATATCTGTCATCATCAATCCCAACCAACACATTTTCTCTCTC 418
 Db 906 ntngday-mvsvngnngnrgnrrhannnarnanandavssnnrnh 948

[illegible]

```

CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC (Q51735). It hybridized to all spp. of mycobacteria tested, but
CC cross reacted to a few non-mycobacterial spp. The probe may
CC be useful as an initial screen for mycobacterial infection.
CC See also Q51735-45 and Q51747-59.
CC Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;
SQ
Query Match 7.6%; Score 35; DB 9; Length 91;
Best Local Similarity 5.7%; Pred. No. 1.18e-03;
Matches 3; Conservative 41; Mismatches 9; Indels 0; Gaps 0;

Ddb 8 gcgssvshyvvvhhshhshhvvhhvhhvvvhhvhhvhhvhhvhhvhhvhhvsv 60
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : :
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 295 GCGGCAAGCGGTCTCTATAATACATTTCCTACATCTCTCTCTCTCTCACA 347

RESULT 7
ID N81164 standard; DNA; 204 BP.
AC N81164;
DT 08-NOV-1990 (first entry)
DE Base substituted E.coli beta-galactosidase alpha-fragment.
KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS Escherichia coli.
SC Key Location/Qualifiers
FF misc_feature 19..69
FT /*tag= a
FT /function= multiple cloning site
FT primer_bind 187..204
FT /*tag= b
FT EP-285123-A.
PN 05-MAY-1988.
PD 30-MAR-1988; 105163.
PF 03-APR-1987; US-034819.
PR (SUSO) SUOMEN SOKERI OY.
PA Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
PI WPI: 88-279927/40.
PT Introducing random point mutations into nucleic acids -
DR by prep of single stranded template, annealing a primer, elongation,
PS misincorporation, completion of molecules and screening.
PS Disclosure; p: English.
CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a pop of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also F80375.
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 7.6%; Score 35; DB 1; Length 204;
Best Local Similarity 6.8%; Pred. No. 1.18e-03;
Matches 6; Conservative 51; Mismatches 30; Indels 1; Gaps 1;

Ddb 99 bnvyrdynrdaaavccyrrsvkdyccynachdhdyvbbvynvhn-hnncnccbn 157
| : : : : : | | | : : : : : : : : : : : : : : : : : : : : : :
| : : : : : | | | : : : : : : : : : : : : : : : : : : : : : :
Cp 434 CGCGGTGAGGAAAAAGAGGAGAGAAAAAGTGTGTTTGGGATTGATTGACAAGT 375

Ddb 158 nhvchnvbnhnrwnayvrhdarddhv 185
: : : : : : : : : : : | : : : : :
Cp 374 ATAATGTTTAAAGAAGAGTGATGGGAT 347

RESULT 8
ID Q51746 standard; CDNA; 91 BP.
AC Q51746;
DT 31-MAY-1994 (first entry)
DE Oligonucleotide probe MK14-A
KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;

```


FT	exon	2941..3213
FT	/*tag= e	
FT	intron	3214..5031
FT	/*tag= f	
FT	exon	5032..6600
FT	/*tag= g	
FT	intron	6601..6933
FT	/*tag= h	
FT	exon	6934..6951
FT	/*tag= i	
PN	W09535024-A1.	
PD	28-DEC-1995.	
PF	16-JUN-1995; PF	U07754.
PR	17-JUN-1994; US-	261663.
PA	(REGC) UNIV CALIFORNIA.	
PA	(USDA) US SEC OF AGRIC.	
PI	Baker BJ Whitham SA;	
DR	WFI; 96-058144/06.	
DR	P-PSDB; R88122.	
PT	Plant virus resistance gene N sequences from tobacco - useful for	
PT	generating transgenic Solanaceous plants resistant to Tobacco Mosaic	
PT	Virus	
PS	Claim 1; Page 52-60; 98pp; English.	
CC	The N gene (r09341) of Nicotiana glutinosa codes for a protein	
CC	(R88123) that mediates resistance to tobacco mosaic virus (TMV).	
CC	The gene was isolated from a genomic library of N. glutinosa by	
CC	screening using a cDNA clone. The N gene can be used to generate	
CC	transgenic plants, esp. Solanaceae, resistant to TMV, or as a	
CC	probe to isolate homologous genes.	
SQ	Sequence 7400 BP; 2495 A; 1147 C; 1246 G; 2512 T;	
Query Match 5.8%; Score 27; DB 17; Length 7400;		
Best Local Similarity 72.9%; Pred. No. 1.00e+00;		
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps		
Db	3645 atatttcgtgaattttttaaaattgtcacaataatgcaaatgaaaaattaaattttt 3703	
Cp	91 ATATTAATAATTAAATTTGAAATATTCAGTTTAATAATTGATATGATAAATTAAATTTT 33	
RESULT 13		
ID	T41852 standard; DNA; 9789 BP.	
DT	T41852;	
DT	20-FEB-1997 (first entry)	
DE	cDNA encoding plasmodium falciapum erythrocyte membrane protein.	
KW	plasmodium falciapum; erythrocyte membrane protein; malaria;	
KW	detection; identification; treatment; prevention; parasite; ss.	
OS	Plasmodium falciapum MC type.	
FS	Key Location/Qualifiers	
FT	CDS 326..9497	
FT	/*tag= a	
ET	/product= Erythrocyte membrane protein	
FT	misc.feature 518..520	
FT	/*tag= b	
FT	/transl_except= GTA encodes Tyrosine	
FT	misc.feature 656..658	
FT	/*tag= c	
FT	/transl_except= ATT encodes Leucine	
FT	misc.feature 2909..2911	
FT	/*tag= d	
FT	/transl_except= AAC encodes Aspartic acid	
FT	misc.feature 3461..3463	
FT	/*tag= e	
FT	/transl_except= GAA encodes Glutamine	
FT	misc.feature 5546..5548	
FT	/*tag= f	
FT	/transl_except= CCT encodes Arginine	
FT	misc.feature 6254..6256	
FT	/*tag= g	
FT	/transl_except= AAT encodes Lysine	
FT	misc.feature 6257..6259	
FT	/*tag= h	
FT	/transl_except= ATA encodes Tyrosine	

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
on: Tue Feb 3 19:49:46 1998; MasPar time 258.59 Seconds
Tabular output not generated.

Release 2.1D John F. Collins, BioComputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

on: Tue Feb 3 19:49:46 1998; MasPar time 258.59 Seconds
Tabular output not generated.

Title: >US-08-713-928A-5
Description: (1-463) from US08713928A.seq
Perfect Score: 463
N.A. Sequence: 1 CAAATACGATATACCGAATA.....CCGGTGAAGTAAAGAGSTC 463
Comp: GTATGCTAATAGGCTTAT.....GGCCACTTTCATTCGTCWAG

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 665703 seqs, 246912890 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

EST-A
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99 100:EST100 101:EST101 102:EST102 103:EST103
104:EST104 105:EST105 106:EST106 107:EST107 108:EST108
109:EST109 110:EST110 111:EST111 112:EST112 113:EST113
114:EST114 115:EST115 116:EST116 117:EST117 118:EST118
119:EST119 120:EST120 121:EST121 122:EST122 123:EST123
124:EST124 125:EST125 126:EST126 127:EST127 128:EST128
129:EST129 130:EST130 131:EST131 132:EST132 133:EST133
134:EST134 135:EST135 136:EST136 137:EST137 138:EST138
139:EST139 140:EST140 141:EST141 142:EST142 143:EST143
144:EST144 145:EST145 146:EST146 147:EST147 148:EST148
149:EST149 150:EST150 151:EST151 152:EST152 153:EST153
154:EST154 155:EST155 156:EST156 157:EST157 158:EST158
159:EST159 160:EST160 161:EST161 162:EST162 163:EST163
164:EST164 165:EST165 166:EST166 167:EST167 168:EST168
169:EST169 170:EST170 171:EST171 172:EST172 173:EST173
174:EST174 175:EST175 176:EST176 177:EST177 178:EST178

Database:

EST-B
99:EST99 100:EST100 101:EST101 102:EST102 103:EST103
104:EST104 105:EST105 106:EST106 107:EST107 108:EST108
109:EST109 110:EST110 111:EST111 112:EST112 113:EST113
114:EST114 115:EST115 116:EST116 117:EST117 118:EST118
119:EST119 120:EST120 121:EST121 122:EST122 123:EST123
124:EST124 125:EST125 126:EST126 127:EST127 128:EST128
129:EST129 130:EST130 131:EST131 132:EST132 133:EST133
134:EST134 135:EST135 136:EST136 137:EST137 138:EST138
139:EST139 140:EST140 141:EST141 142:EST142 143:EST143
144:EST144 145:EST145 146:EST146 147:EST147 148:EST148
149:EST149 150:EST150 151:EST151 152:EST152 153:EST153
154:EST154 155:EST155 156:EST156 157:EST157 158:EST158
159:EST159 160:EST160 161:EST161 162:EST162 163:EST163
164:EST164 165:EST165 166:EST166 167:EST167 168:EST168
169:EST169 170:EST170 171:EST171 172:EST172 173:EST173
174:EST174 175:EST175 176:EST176 177:EST177 178:EST178

179:EST179 180:EST180 181:EST181 182:EST182 183:EST183
184:EST184 185:EST185 186:EST186 187:EST187 188:EST188
189:EST189 190:EST190 191:EST191 192:EST192 193:EST193
194:EST194 195:EST195 196:EST196

Statistics: Mean 10.824; Variance 4.238; scale 2.554

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
C 1	30	6.5	418 183	N51049	YV30d11.s1 Soares fet	5.13e-05
C 2	28	6.0	365 26	R44938	YQ32h08.s1 Homo sapie	1.40e-03
C 3	28	6.0	479 124	W08634	NB42b12.r1 Soares mou	1.40e-03
C 4	27	5.8	175 153	AA072359	mf13b08.r1 Life Tech	6.98e-03
C 5	27	5.8	205 170	AA153231	mn31b12.r1 Beddington	6.98e-03
C 6	27	5.8	265 17	T18158	0742c3 Plasmodium fal	6.98e-03
C 7	27	5.8	405 192	AA135806	zn93b08.s1 Stratagene	6.98e-03
C 8	26	5.6	236 146	AA117266	mn22f11.r1 Beddington	3.36e-02
C 9	26	5.6	255 99	N95064	ZB32d06.s1 Homo sapie	3.36e-02
C 10	26	5.6	288 145	AA104461	mo55b01.r1 Life Tech	3.36e-02
C 11	26	5.6	290 4	T67279	Y453g12.r4 Homo sapie	3.36e-02
C 12	26	5.6	356 14	R06872	Yf11f11.s1 Homo sapie	3.36e-02
C 13	26	5.6	361 188	AA034637	mh17a06.r1 Soares mou	3.36e-02
C 14	26	5.6	371 51	R83527	Yp15h02.r1 Homo sapie	3.36e-02
C 15	26	5.6	375 176	AA184881	mu46a04.r1 Soares mou	3.36e-02
C 16	26	5.6	377 183	AA214473	Zq92d05.r1 Stratagene	3.36e-02
C 17	26	5.6	407 168	AA099054	zn45d11.r1 Stratagene	3.36e-02
C 18	26	5.6	444 85	N59340	Yz86d10.s1 Homo sapie	3.36e-02
C 19	26	5.6	485 76	N27300	Yw71f10.s1 Homo sapie	3.36e-02
C 20	26	5.6	507 159	W82081	ne96h06.r1 Soares mou	3.36e-02
C 21	26	5.6	516 192	AA141823	CK02516.3prime Drosop	3.36e-02
C 22	26	5.6	573 57	T23157	C2H2-218T Homo sapien	3.36e-02
C 23	26	5.6	579 193	AA143490	Zo65h12.s1 Stratagene	3.36e-02
C 24	25	5.4	142 146	AA117240	mn19f12.r1 Beddington	1.56e-01
C 25	25	5.4	191 15	T02709	0064M3 Plasmodium fal	1.56e-01
C 26	25	5.4	202 95	AFTS1181	A. thaliana transcrib	1.56e-01
C 27	25	5.4	204 26	R42498	Yg02h06.s1 Homo sapie	1.56e-01
C 28	25	5.4	219 176	AA184945	mt88c02.r1 Soares mou	1.56e-01
C 29	25	5.4	226 93	H82133	Ys71h09.r1 Homo sapie	1.56e-01
C 30	25	5.4	233 33	R65859	Y123a08.r1 Homo sapie	1.56e-01
C 31	25	5.4	274 150	N97347	0133k3 gmbpFHB3.1 G.	1.56e-01
C 32	25	5.4	277 16	T09966	0328m7 Plasmodium fal	1.56e-01
C 33	25	5.4	311 119	HSCOHG121	H. sapiens partial cd	1.56e-01
C 34	25	5.4	311 150	N98059	2144C3 czapFDd2.1, D	1.56e-01
C 35	25	5.4	336 15	T02659	0027M7 Plasmodium fal	1.56e-01
C 36	25	5.4	354 176	AA185015	mt88g11.r1 Soares mou	1.56e-01
C 37	25	5.4	358 153	W45543	Zc26c01.s1 Soares sen	1.56e-01
C 38	25	5.4	388 74	N21750	SMNHAD017043SK Schis	1.56e-01
C 39	25	5.4	410 130	W46716	md20e05.r1 Life Tech	1.56e-01
C 40	25	5.4	449 130	W58509	ZM20e01.s1 Soares fet	1.56e-01
C 41	25	5.4	458 188	AA032099	SWAMCA1090SK Brugla m	1.56e-01
C 42	25	5.4	463 172	AA164955	Zq42g07.s1 Stratagene	1.56e-01
C 43	25	5.4	496 153	W49719	Zc43f11.s1 Soares sen	1.56e-01
C 44	25	5.4	504 40	H15474	Ym29b05.r1 Homo sapie	1.56e-01
C 45	25	5.4	515 30	R59710	Yh1ld03.s1 Homo sapie	1.56e-01

ALIGNMENTS

RESULT 1 N51049 418 bp mRNA 28-JAN-1997
LOCUS YV30d11.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
DEFINITION 244245 3' similar to contains Alu repetitive element; contains L1.t2
L1 repetitive element ;
ACCESSION N51049
NID 91192215
KEYWORDS EST.

```

SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
AUTHORS     Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 418)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevaskis, E.,
Waterson, R., Williamson, A., Wohlmann, P. and Wilson, R.
washington.wustl.edu
WashU-Merck EST Project
Unpublished (1995)

TITLE      Contact: Wilson RK
JOURNAL    Washington University School of Medicine
COMMENT    4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 716 Std Error: 0.00
Seq primer: ml3 -40 forward
High quality sequence stop: 337.
Location/Qualifiers
FEATURES   source
1..418
/organism="Homo sapiens"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACGGAGAAATTAATAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
/clone="244245"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
complement(<1..>418)
BASE COUNT 106 a 67 c 87 g 158 t
ORIGIN
mRNA
Query Match 6.5%; Score 30; DB 183; Length 418;
Best Local Similarity 69.2%; Pred. No. 5.13e-05;
Matches 63; Conservative 0; Mismatches 27; Indels 1; Gaps 1;
206 ttgtttttgaaatttgaatggaa-attgtatttttttcattatgattgtttttt 264
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cp 286 TTTTTCATTTTATATTTTGAATGGAGTATATCTTTTACAATAATATCTTCTTTT 227
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 265 ttttttttttgagatgaggtcttgtttgt 295
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cp 226 TTTAATTTTGAAGTAAAGTTTATATAGT 196
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
LOCUS      R44938
DEFINITION Yg32h08.s1 Homo sapiens cDNA clone 34206 3'.
ACCESSION R44938
NID        9823205
KEYWORDS   EST.
SOURCE     human clone=34206 library=Soares infant brain INIB vector=Lafmid BA
host=DH10B (ampicillin resistant) primer=Promega -21ml3 Rsite1=Not
I Rsite2=Hind III whole brain from a 73 days post natal female. 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
AAGTGGAGATTCGCGCGCAGAGATTTTGTGTTTGT 3']; double-stranded
cDNA was ligated to Hind III adaptors (Pharmacia), digested with
Not I and directionally cloned into the Not I and Hind III sites of
the Lafmid BA vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.Fatima
Bonaldo.

```

```

Bonaldo.
Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrápoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 365)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)

GDB: G00-406-553
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
High quality sequence stops: 240
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Location/Qualifiers
FEATURES   source
1..365
/organism="Homo sapiens"
/clone="34206"
BASE COUNT 109 a 56 c 59 g 130 t 11 others
ORIGIN
Query Match 6.0%; Score 28; DB 26; Length 365;
Best Local Similarity 83.3%; Pred. No. 1.40e-03;
Matches 35; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Db 11 ttttttttttttaactttttcaagttaattgttatgtacagt 52
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cp 237 TTTCTTTTTTTTAAATTTTTCGAAGTAAAGTTTATATAGT 196
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
LOCUS      W08634
DEFINITION mb42b12.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 332063
5'.
ACCESSION W08634
NID        g1282657
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus.
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE 1 (bases 1 to 479)
AUTHORS   Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

```

[illegible]

BASE COUNT	160 a	11 c	11 g	23 t
ORIGIN				
Query Match	5.8%;	Score 27;	DB 170;	Length 205;
Best Local Similarity	65.2%;	Pred. No.	6,98e-03;	
Matches	58;	Conservative	0;	Mismatches 31; Indels 0; Gaps 0;
Db	56	ataataacacctttctcacaaaaaa	aaaaaaaaaaaaaaaaaaaaaaaaaaa	115
-				
QY	199	ATAAATAAACTTTACTTCAAAAAAT	TAAAAAAAAGGAAGTATATTATTG	TAAGAATAA 258
Db	116	aaaaa	aaaaaaaaaaaaaaaaaaaaa	144
-				
QY	259	TACTCCATTCAAATATAAATGAAAAA	A 287	
RESULT	6			
LOCUS	Tl18158	265 bp	mRNA	EST 30-AUG-1994
DEFINITION	O742c3 Plasmodium falciparum	cdna clone 0742c 5'.		
ACCESSION	Tl18158			
KEYWORDS	G462944			
SOURCE	EST.			
DESCRIPTION	malaria parasite clone-0742c library-czapfEDd2.1, Debopam Chakrabarti strain-Dd2 vector-Lambda ZAP II host=E. coli XL-1 blue primer-T3 Raitel-EcoR I Raited-Xho I PolyA+ RNA, from asynchronous blood stage parasites of the Dd2 isolate cultured in vitro, was reverse transcribed using an oligo dt-xho i primer. Second strand was prepared using RNase H and DNA polymerase I. EcoR I adapters were ligated to the cdna, and it was digested with Xho I. Prepared fragments were ligated into EcoR I + Xho I digested lambda ZAP II vector..			
ORGANISM	Plasmodium falciparum			
REFERENCE	Eucaryotae; Protozoa; Apicomplexa; Sporozoa; Coccidia; Eucoccidiidae; Haemosporina; Plasmodium.			
AUTHORS	1 (bases 1 to 265)			
TITLE	Chakrabarti,D., Reddy,G.R., Dame,J.B., Almira,E.C., Laipis,P.J., Ferl,R.J., Yang,T.P., Rowe,T.C. and Schuster,S.M.			
JOURNAL	Analysis of Expressed Sequence Tags from Plasmodium Falciparum Mol. Biochem. Parasitol. 66, 97-104 (1994)			
MEDLINE	95075403			
COMMENT	Contact: Debopam Chakrabarti Biotechnology & Infectious Diseases University of Florida Box 110880, Gainesville, FL 32611 Tel.: 9043924700 ext. 5817 Fax: 9043929704 Email: dchake@icbr.ifas.ufl.edu. Location/Qualifiers 1..265 /organism="Plasmodium falciparum" /clone="0742c" /strain="dd2"			
NATURES				
source				
BASE COUNT	107 a	16 c	25 g	111 t
ORIGIN				
Query Match	5.8%;	Score 27;	DB 17;	Length 265;
Best Local Similarity	65.9%;	Pred. No.	6,98e-03;	
Matches	56;	Conservative	0;	Mismatches 29; Indels 0; Gaps 0;
Db	12	tctatgttaattggaagtatatataaatatataatattatgatataatattt	71	
-				
Cp	283	TTCATTTTATATTTTGATGTGAGTATTCTTC	TTCATAATAATATCTTCTTTTTTTT	224
Db	72	attttttgcacacatattttacat	96	
-				
Cp	223	AATTTTGAAGTAAGTTTATAT	199	
RESULT	7			
LOCUS	AA135806	405 bp	mRNA	EST 30-NOV-1996
DEFINITION	zn93808.s1 Stratiagene lung carcinoma 937218 Homo sapiens cdna clone 565719 3' similar to contains Alu repetitive element;contains			

BASE COUNT

[illegible]

FEATURES
source

MRNA	BASE COUNT	ORIGIN
1	100	1
2	100	2
3	100	3
4	100	4
5	100	5
6	100	6
7	100	7
8	100	8
9	100	9
10	100	10
11	100	11
12	100	12
13	100	13
14	100	14
15	100	15
16	100	16
17	100	17
18	100	18
19	100	19
20	100	20
21	100	21
22	100	22
23	100	23
24	100	24
25	100	25
26	100	26
27	100	27
28	100	28
29	100	29
30	100	30
31	100	31
32	100	32
33	100	33
34	100	34
35	100	35
36	100	36
37	100	37
38	100	38
39	100	39
40	100	40
41	100	41
42	100	42
43	100	43
44	100	44
45	100	45
46	100	46
47	100	47
48	100	48
49	100	49
50	100	50
51	100	51
52	100	52
53	100	53
54	100	54
55	100	55
56	100	56
57	100	57
58	100	58
59	100	59
60	100	60
61	100	61
62	100	62
63	100	63
64	100	64
65	100	65
66	100	66
67	100	67
68	100	68
69	100	69
70	100	70
71	100	71
72	100	72
73	100	73
74	100	74
75	100	75
76	100	76
77	100	77
78	100	78
79	100	79
80	100	80
81	100	81
82	100	82
83	100	83
84	100	84
85	100	85
86	100	86
87	100	87
88	100	88
89	100	89
90	100	90
91	100	91
92	100	92
93	100	93
94	100	94
95	100	95
96	100	96
97	100	97
98	100	98
99	100	99
100	100	100

Search completed: Tue Feb 3 20:05:54 1998
Job time : 968 secs.

W P S R E H (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
n: Tue Feb 3 15:32:19 1998; MasPar time 2.99 Seconds
Tabular output not generated. 81.629 Million cell updates/sec

Title: >US-08-713-928A-10
Description: (1-8) from US08713928A.pep
Perfect Score: 56
Sequence: 1 DYKDDDDK 8

Scoring table: PAM 150
Gap 15

Searched: 95051 seqs, 30469580 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir53
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8
13:unann9 14:unann10 15:unann11 16:unann12 17:unann
18:unrev

Statistics: Mean 21.667; Variance 32.486; scale 0.667

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	48	85.7	145	8	JX0313 mite allergen mag 29	1.73e+01
2	48	85.7	423	14	S59425 SPR28 protein - yeas	1.73e+01
3	48	85.7	836	15	S34399 dynamin 3 - fruit fl	1.73e+01
4	48	85.7	836	15	S15413 dynamin-like protein	1.73e+01
5	48	85.7	836	15	S17974 dynamin-like protein	1.73e+01
6	48	85.7	848	16	I55498 testicular dynamin -	1.73e+01
7	48	85.7	851	15	S11508 p100 protein - rat	1.73e+01
8	48	85.7	851	15	B40671 dynamin, internal fo	1.73e+01
9	48	85.7	864	15	A40671 dynamin, internal fo	1.73e+01
10	48	85.7	883	15	S19795 dynamin-like protein	1.73e+01
11	48	85.7	883	15	S16130 dynamin 4 - fruit fl	1.73e+01
12	48	85.7	1277	13	S54451 hypothetical protein	1.73e+01
13	47	83.9	403	14	S63130 hypothetical protein	2.57e+01
14	47	83.9	645	8	S41372 heat shock protein 7	2.57e+01
15	47	83.9	646	8	A44985 heat shock protein 7	2.57e+01
16	47	83.9	866	15	JC4305 dynamin II - human	2.57e+01
17	47	83.9	868	16	A36878 dynamin 2 - rat	2.57e+01
18	47	83.9	870	16	B53165 dynamin II isoform b	2.57e+01
19	47	83.9	870	16	A53165 dynamin II isoform a	2.57e+01
20	46	82.1	147	4	B46315 E4 protein - human p	3.80e+01

21	46	82.1	289	13	S16726	methyviologen-reduc	3.80e+01
22	46	82.1	306	13	S32834	methyviologen-reduc	3.80e+01
23	45	80.4	199	12	D64482	hypothetical protein	5.59e+01
24	45	80.4	201	9	F40781	ORF6 protein - Autog	5.59e+01
25	45	80.4	201	9	A34146	hypothetical protein	5.59e+01
26	45	80.4	406	14	S59296	probable finger prot	5.59e+01
27	45	80.4	483	13	S46124	hypothetical protein	5.59e+01
28	45	80.4	483	14	JG6196	nuclear protein Enpl	5.59e+01
29	45	80.4	728	12	H64346	hypothetical protein	5.59e+01
30	45	80.4	1088	1	PRTGA	platelet-derived gro	5.59e+01
31	45	80.4	1089	5	S33727	platelet-derived gro	5.59e+01
32	45	80.4	1089	1	PFHUGA	platelet-derived gro	5.59e+01
33	44	78.6	37	18	S21665	S-layer protein - Ac	8.18e+01
34	44	78.6	153	10	B64157	hypothetical protein	8.18e+01
35	44	78.6	308	12	D30315	hypothetical protein	8.18e+01
36	44	78.6	488	4	VCCV4C	coat protein - cauli	8.18e+01
37	44	78.6	488	9	JN0496	hypothetical 56.5K p	8.18e+01
38	44	78.6	489	4	VCCV	coat protein - cauli	8.18e+01
39	44	78.6	490	4	VCCV3	coat protein - cauli	8.18e+01
40	44	78.6	514	8	S18449	variant surface glyc	8.18e+01
41	44	78.6	607	18	S50658	legumin - Gnetum gne	8.18e+01
42	44	78.6	762	10	A34355	cell surface protein	8.18e+01
43	44	78.6	762	13	S44560	alpha,alpha-trehalas	8.18e+01
44	44	78.6	917	13	S07183	hypothetical protein	8.18e+01
45	43	76.8	2206	14	JC5280	voltage-dependent ca	1.19e+02

ALIGNMENTS

RESULT 1
ENTRY JX0313 #type complete
TITLE mite allergen mag 29 protein - house-dust mite (Dermatophagoides farinae)
ORGANISM #formal_name Dermatophagoides farinae
DATE 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 28-Feb-1997
ACCESSIONS JX0313; PC2125
REFERENCE JX0313
#authors Aki, T.; Fujikawa, A.; Wada, T.; Jyo, T.; Shigeta, S.; Murooka, Y.; Oka, S.; Ono, K.
#journal J. Biochem. (1994) 115:435-440
#title Cloning and expression of cDNA coding for a new allergen from the house dust mite, Dermatophagoides farinae: homology with human heat shock cognate proteins in the heat shock protein 70 family.
#accession JX0313
##molecule_type mRNA
##residues 1-145 #label AK1
#accession PC2125
##molecule_type protein
##residues 1-17 #label AK2
##cross-references DBJ:D17676
##note the nucleotide sequence for this amino acid sequence is inconsistent with that for D17676 in having an additional nucleotide A at the code for 141-ile

GENETICS mag29
#gene #superfamily heat shock protein 70
CLASSIFICATION ATP; heat shock; stress-induced protein
KEYWORDS #length 145 #molecular-weight 15594 #checksum 1862
SUMMARY

Query Match 85.7%; Score 48; DB 8; Length 145;
Best Local Similarity 71.4%; Pred. No. 1.73e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 14 ykdedk 20
|||:||||
Qy 2 YKDDDDK 8

RESULT 2 S59425 #type complete
ENTRY SPR28 protein - yeast (Saccharomyces cerevisiae)
TITLE

```
ALTERNATE_NAMES  protein YD9934.03c: protein YDR218c
ORGANISM          #formal_name Saccharomyces cerevisiae
DATE             30-Nov-1995 #sequence_revision 16-Feb-1996 #text_change
                23-Aug-1996
ACCESSIONS       S59425
REFERENCE        S59423
#authors         Murphy, L.; Harris, D.
#submission      submitted to the EMBL Data Library, March 1995
#accession       S59425
#molecule_type  DNA
#residues        1-423 ##label MUR
#cross-references EMBL:Z48612
#experimental_source strain AB972
GENETICS
#gene            SPR28
#map_position    4R
SUMMARY          #length 423 #molecular-weight 48193 #checksum 7812
Query Match      85.7%; Score 48; DB 14; Length 423;
Best Local Similarity 87.5%; Pred. No. 1.73e+01;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 62 dyddddd 69
|||
QY 1 YKDDDDK 8

RESULT 3
ENTRY   #type complete
TITLE   dynamin 3 - fruit fly (Drosophila melanogaster)
ORGANISM
DATE    30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
ACCESSIONS S34399
REFERENCE S34399; S15497
#authors   Chen, M.S.; Obar, R.A.; Schroeder, C.C.; Austin, T.W.;
            Poody, C.A.; Wadsworth, S.C.; Vallee, R.B.
#journal   Nature (1991) 351:583-586
#title     Multiple forms of dynamin are encoded by shibire, a
            Drosophila gene involved in endocytosis.
#cross-references MUID:91260878
#accession S34399
#molecule_type mRNA
#residues  1-836 ##label CHE
#cross-references EMBL:X59448
GENETICS
#gene      shibire
#cross-references FlyBase:FBgn0003392
#keywords  alternative splicing; microtubule binding; P-loop
#region    nucleotide-binding motif A (P-loop)\
127-132   #region nucleotide-binding motif B\
200-203   #region nucleotide-binding #status predicted
SUMMARY   #length 836 #molecular-weight 93758 #checksum 9216
Query Match      85.7%; Score 48; DB 15; Length 836;
Best Local Similarity 71.4%; Pred. No. 1.73e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 547 ykdedek 553
|||
QY 2 YKDDDDK 8

RESULT 4
ENTRY   #type complete
TITLE   dynamin-like protein - fruit fly (Drosophila sp.)
ORGANISM #formal_name Drosophila sp.
DATE     21-Nov-1993 #sequence_revision 03-Nov-1995 #text_change
ACCESSIONS S15413
REFERENCE S15413
#authors   van der Bliek, A.M.; Meyerowitz, E.M.
#cross-references FlyBase:FBgn0003392
#keywords  alternative splicing; microtubule binding; P-loop
#region    nucleotide-binding motif A (P-loop)\
127-132   #region nucleotide-binding motif B\
200-203   #region nucleotide-binding #status predicted
SUMMARY   #length 836 #molecular-weight 93758 #checksum 9216
Query Match      85.7%; Score 48; DB 15; Length 836;
Best Local Similarity 71.4%; Pred. No. 1.73e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 547 ykdedek 553
|||
QY 2 YKDDDDK 8

ALTERNATE_NAMES  protein YD9934.03c: protein YDR218c
ORGANISM          #formal_name Saccharomyces cerevisiae
DATE             30-Nov-1995 #sequence_revision 16-Feb-1996 #text_change
                23-Aug-1996
ACCESSIONS       S59425
REFERENCE        S59423
#authors         Murphy, L.; Harris, D.
#submission      submitted to the EMBL Data Library, March 1995
#accession       S59425
#molecule_type  DNA
#residues        1-423 ##label MUR
#cross-references EMBL:Z48612
#experimental_source strain AB972
GENETICS
#gene            SPR28
#map_position    4R
SUMMARY          #length 423 #molecular-weight 48193 #checksum 7812
Query Match      85.7%; Score 48; DB 14; Length 423;
Best Local Similarity 87.5%; Pred. No. 1.73e+01;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 62 dyddddd 69
|||
QY 1 YKDDDDK 8

RESULT 3
ENTRY   #type complete
TITLE   dynamin 3 - fruit fly (Drosophila melanogaster)
ORGANISM
DATE    30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
ACCESSIONS S34399
REFERENCE S34399; S15497
#authors   Chen, M.S.; Obar, R.A.; Schroeder, C.C.; Austin, T.W.;
            Poody, C.A.; Wadsworth, S.C.; Vallee, R.B.
#journal   Nature (1991) 351:583-586
#title     Multiple forms of dynamin are encoded by shibire, a
            Drosophila gene involved in endocytosis.
#cross-references MUID:91260878
#accession S34399
#molecule_type mRNA
#residues  1-836 ##label CHE
#cross-references EMBL:X59448
GENETICS
#gene      shibire
#cross-references FlyBase:FBgn0003392
#keywords  alternative splicing; microtubule binding; P-loop
#region    nucleotide-binding motif A (P-loop)\
127-132   #region nucleotide-binding motif B\
200-203   #region nucleotide-binding #status predicted
SUMMARY   #length 836 #molecular-weight 93758 #checksum 9216
Query Match      85.7%; Score 48; DB 15; Length 836;
Best Local Similarity 71.4%; Pred. No. 1.73e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 547 ykdedek 553
|||
QY 2 YKDDDDK 8

RESULT 4
ENTRY   #type complete
TITLE   dynamin-like protein - fruit fly (Drosophila sp.)
ORGANISM #formal_name Drosophila sp.
DATE     21-Nov-1993 #sequence_revision 03-Nov-1995 #text_change
ACCESSIONS S15413
REFERENCE S15413
#authors   van der Bliek, A.M.; Meyerowitz, E.M.
#cross-references FlyBase:FBgn0003392
#keywords  alternative splicing; microtubule binding; P-loop
#region    nucleotide-binding motif A (P-loop)\
127-132   #region nucleotide-binding motif B\
200-203   #region nucleotide-binding #status predicted
SUMMARY   #length 836 #molecular-weight 93758 #checksum 9216
Query Match      85.7%; Score 48; DB 15; Length 836;
Best Local Similarity 71.4%; Pred. No. 1.73e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 547 ykdedek 553
|||
QY 2 YKDDDDK 8
```

```
#journal      Nature (1991) 351:411-414
#title        Dynamn-like protein encoded by the Drosophila shibire gene
              associated with vesicular traffic.
#cross-references MUID:91238973
#accession    S15413
#status       preliminary; nucleic acid sequence not shown
#molecule_type mRNA
#residues     1-836 ##label NAT
#cross-references EMBL:X59435
GENETICS
#gene         shibire
#cross-references FlyBase:FBgn0003392
SUMMARY       #length 836 #molecular-weight 93671 #checksum 9040
Query Match   85.7%; Score 48; DB 15; Length 836;
Best Local Similarity 71.4%; Pred. No. 1.73e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 547 ykdedek 553
|||
QY 2 YKDDDDK 8

RESULT 5
ENTRY   #type complete
TITLE   dynamn-like protein - fruit fly (Drosophila melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE     13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change
              16-Feb-1997
ACCESSIONS S17974
REFERENCE S17974
#authors   van der Bliek, A.M.
#submission submitted to the EMBL Data Library, May 1991
#accession S17974
#status     preliminary
#molecule_type mRNA
#residues   1-836 ##label VAN
#cross-references EMBL:X59435
GENETICS
#gene      FlyBase:shi
#cross-references FlyBase:FBgn0003392
SUMMARY   #length 836 #molecular-weight 93730 #checksum 9048
Query Match   85.7%; Score 48; DB 15; Length 836;
Best Local Similarity 71.4%; Pred. No. 1.73e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 547 ykdedek 553
|||
QY 2 YKDDDDK 8

RESULT 6
ENTRY   #type complete
TITLE   testicular dynamin - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE     26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change
              26-Jul-1996
ACCESSIONS I55498
REFERENCE I55498
#authors   Nakata, T.; Takemura, R.; Hirokawa, N.
#journal   J. Cell Sci. (1993) 105:1-5
#title     A novel member of the dynamin family of GTP-binding proteins
              is expressed specifically in the testis.
#cross-references MUID:93366923
#accession I55498
#status     preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues   1-848 ##label RES
#cross-references GB:D14076; NID:g391871; CDS_PID:g391872
SUMMARY   #length 848 #molecular-weight 95595 #checksum 9171
Query Match   85.7%; Score 48; DB 16; Length 848;
```

Best Local Similarity 71.4%; Pred. No. 1.73e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 549 ykdddek 555
||||:|
QY 2 YKDDDDK 8

RESULT 7
ENTRY S11508 #type complete
TITLE D100 protein - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995
ACCESSIONS S11508
REFERENCE S11508
#authors Obar, R.A.; Collins, C.A.; Hammarback, J.A.; Shpetner, H.S.;
Vallee, R.B.
#journal Nature (1990) 347:256-261
#title Molecular cloning of the microtubule-associated
mechanochemical enzyme dynamin reveals homology with a new
family of GTP-binding proteins.
#cross-references MUID:90384564
#accession S11508
#molecule_type mRNA
#residues 1-851 #label OBA
#cross-references EMBL:X54531
SUMMARY #length 851 #molecular-weight 95927 #checksum 7562

Query Match 85.7%; Score 48; DB 16; Length 851;
Best Local Similarity 71.4%; Pred. No. 1.73e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 553 ykdddek 559
||||:|
QY 2 YKDDDDK 8

RESULT 8
ENTRY B40671 #type complete
TITLE dynamin, internal form 2, short C-terminal form - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 03-May-1994 #sequence_revision 03-May-1994 #text_change 03-May-1994
ACCESSIONS B40671
REFERENCE B40671
#authors van der Bliek, A.M.; Redelmeier, T.E.; Damke, H.; Tisdale,
E.J.; Meyerowitz, E.M.; Schmid, S.L.
#journal J. Cell Biol. (1993) 122:553-563
#title Mutations in human dynamin block an intermediate stage in
coated vesicle formation.
#accession B40671
#status preliminary
#molecule_type mRNA
#residues 1-851 #label VAN
#cross-references GB:L07807
KEYWORDS alternative splicing; GTP binding
SUMMARY #length 851 #molecular-weight 96039 #checksum 7509

Query Match 85.7%; Score 48; DB 15; Length 851;
Best Local Similarity 71.4%; Pred. No. 1.73e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 553 ykdddek 559
||||:|
QY 2 YKDDDDK 8

RESULT 9
ENTRY A40671 #type complete
TITLE dynamin, internal form 1, long C-terminal form - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 03-May-1994 #sequence_revision 03-May-1994 #text_change 03-May-1994

03-May-1994
ACCESSIONS A40671
REFERENCE A40671
#authors van der Bliek, A.M.; Redelmeier, T.E.; Damke, H.; Tisdale,
E.J.; Meyerowitz, E.M.; Schmid, S.L.
#journal J. Cell Biol. (1993) 122:553-563
#title Mutations in human dynamin block an intermediate stage in
coated vesicle formation.
#accession A40671
#status preliminary
#molecule_type mRNA
#residues 1-864 #label VAN
#cross-references GB:L07807
KEYWORDS alternative splicing; GTP binding
SUMMARY #length 864 #molecular-weight 97261 #checksum 8496

Query Match 85.7%; Score 48; DB 15; Length 864;
Best Local Similarity 71.4%; Pred. No. 1.73e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 553 ykdddek 559
||||:|
QY 2 YKDDDDK 8

RESULT 10
ENTRY S17975 #type complete
TITLE dynamin-like protein - fruit fly (Drosophila sp.)
ORGANISM #formal_name Drosophila sp.
DATE 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 16-Feb-1997
ACCESSIONS S17975
REFERENCE S15413
#authors van der Bliek, A.M.; Meyerowitz, E.M.
#journal Nature (1991) 351:411-414
#title Dynamin-like protein encoded by the Drosophila shibire gene
associated with vesicular traffic.
#cross-references MUID:91238973
#accession S17975
#status preliminary; nucleic acid sequence not shown
#molecule_type mRNA
#residues 1-883 #label BLI
#cross-references EMBL:X59435
GENETICS
#gene shibire
#introns 834/3
SUMMARY #length 883 #molecular-weight 98508 #checksum 580

Query Match 85.7%; Score 48; DB 15; Length 883;
Best Local Similarity 71.4%; Pred. No. 1.73e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 547 ykdddek 553
||||:|
QY 2 YKDDDDK 8

RESULT 11
ENTRY S16130 #type complete
TITLE dynamin 4 - fruit fly (Drosophila melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 28-Feb-1997
ACCESSIONS S16130; S15498
REFERENCE S16130
#authors Chen, M.S.; Obar, R.A.; Schroeder, C.C.; Austin, T.W.;
Poody, C.A.; Wadsworth, S.C.; Vallee, R.B.
#journal Nature (1991) 351:583-586
#title Multiple forms of dynamin are encoded by shibire, a
Drosophila gene involved in endocytosis.
#cross-references MUID:91260878
#accession S16130

```

##molecule_type mRNA
##residues 1-883 #label WAD
##cross-references EMBL:X59449
GENETICS
#gene shibire
##cross-references FlyBase:FBgn0003392
FEATURE
- 33-40 #region nucleotide-binding motif A (P-loop)\
- 127-132 #region nucleotide-binding motif B\
-200-203 #region nucleotide-binding #status predicted
SUMMARY
#length 883 #molecular-weight 98536 #checksum 748
Query Match 85.7%; Score 48; DB 15; Length 883;
Best Local Similarity 71.4%; Pred. No. 1.73e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 547 ykdedek 553
QY 2 YKDDDDK 8
RESULT 12
ENTRY
TITLE S54451 #type complete
#molecule_type DNA
#formal_name hypothetical protein YNR076c - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES YNR076c
ORGANISM Saccharomyces cerevisiae
DATE 08-Jul-1995 #sequence_revision 09-Mar-1996 #text_change
ACCESSIONS S54451; S52836
REFERENCE S54451
#authors Gentiles, S.; Bowman, S.
#submission submitted to the EMBL Data Library, May 1995
#accession S54451
#molecule_type DNA
##residues 1-682 #label GEN
##cross-references EMBL:Z49259
##experimental_source strain AB972
REFERENCE S52814
#authors Pearson, D.; Bowman, S.
#submission submitted to the EMBL Data Library, April 1995
#accession S52836
#molecule_type DNA
##residues 659-1277 #label PEA
##cross-references EMBL:Z48952
##experimental_source strain AB972
GENETICS
#map_position 13R
SUMMARY
#length 1277 #molecular-weight 147040 #checksum 8885
Query Match 85.7%; Score 48; DB 13; Length 1277;
Best Local Similarity 85.7%; Pred. No. 1.73e+01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 1265 ykdedek 1271
QY 1 YKDDDD 7
RESULT 13
ENTRY
TITLE S63130 #type complete
#molecule_type DNA
#formal_name hypothetical protein YNL175c - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES YNL175c
ORGANISM Saccharomyces cerevisiae
DATE 27-Apr-1996 #sequence_revision 03-May-1996 #text_change
ACCESSIONS S63130
REFERENCE S63122
#authors Obermaier, B.; Piravandi, E.; Rinke, M.; Domdey, H.
#submission submitted to the Protein Sequence Database, April 1996

```

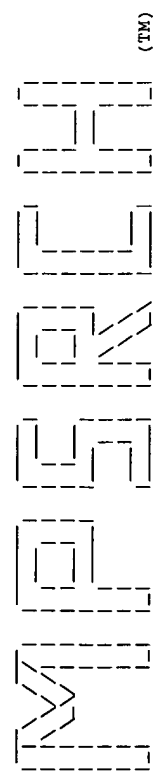
```

#accession S63130
#molecule_type DNA
##residues 1-403 #label OBE
##cross-references EMBL:Z71451
##experimental_source strain S288C
GENETICS
#map_position 14L
CLASSIFICATION #superfamily ribonucleoprotein repeat homology
SUMMARY #length 403 #molecular-weight 45667 #checksum 9240
Query Match 83.9%; Score 47; DB 14; Length 403;
Best Local Similarity 50.0%; Pred. No. 2.57e+01;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 79 eykedaek 86
QY 1 DYKDDDDK 8
RESULT 14
ENTRY
TITLE S41372 #type complete
#molecule_type DNA
#formal_name heat shock protein - yeast (Hansenula polymorpha)
ORGANISM Hansenula polymorpha
DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
ACCESSIONS S41372
REFERENCE S41372
#authors Diesel, A.A.; Roggenkamp, R.R.
#submission submitted to the EMBL Data Library, January 1994
#description HSP70 genes of the yeast Hansenula polymorpha.
#accession S41372
#status preliminary
#molecule_type DNA
##residues 1-645 #label DIE
##cross-references EMBL:Z29379
CLASSIFICATION #superfamily heat shock protein 70
KEYWORDS ATP
SUMMARY #length 645 #molecular-weight 70137 #checksum 6864
Query Match 83.9%; Score 47; DB 8; Length 645;
Best Local Similarity 57.1%; Pred. No. 2.57e+01;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 525 ykedaek 531
QY 2 YKDDDDK 8
RESULT 15
ENTRY
TITLE A44985 #type complete
#molecule_type DNA
#formal_name heat shock protein 70.1 - Theileria annulata
ORGANISM Theileria annulata
DATE 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change
ACCESSIONS A44985
REFERENCE A44985
#authors Mason, P.J.; Shields, B.R.; Tait, A.; Beck, P.; Hall, R.
#journal Mol. Biochem. Parasitol. (1989) 37:27-36
#title Sequence and expression of a gene from Theileria annulata coding for a 70-kilodalton heat-shock protein.
#accession A44985
#status preliminary
#molecule_type DNA
##residues 1-646 #label MAS
##cross-references GB:J04653
CLASSIFICATION #superfamily heat shock protein 70
KEYWORDS ATP
SUMMARY #length 646 #molecular-weight 70991 #checksum 5028
Query Match 83.9%; Score 47; DB 8; Length 646;
Best Local Similarity 57.1%; Pred. No. 2.57e+01;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

Db 526 ykeedek 532
11:11:1
QY 2 YKDDDK 8

Search completed: Tue Feb 3 15:32:39 1998
Job time : 20 secs.



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run: Tue Feb 3 15:31:48 1998; MasPar time 2.04 Seconds
Tabular output not generated. 83.171 Million cell updates/sec

Title: >US-08-713-928A-10
Description: (1-8) from US08713928A.pep
Perfect Score: 56
Sequence: 1 DYKDDDDK 8

Scoring table: PAM 150
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot34
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 22.506; Variance 27.580; scale 0.816

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Rank	No.	Score	Match	Length	ID	Description	Pred. No.
1	48	85.7	145	6	MA29_DERFA	ALLERGEN MAG29 (FRAGM	4.53e+00
2	48	85.7	848	3	DYN3_RAT	DYNAMIN 3 (DYNAMIN T	4.53e+00
3	48	85.7	851	3	DYN1_RAT	DYNAMIN-1 (D100) (DYN	4.53e+00
4	48	85.7	861	3	DYN1_MOUSE	DYNAMIN-1 (DYNAMIN BR	4.53e+00
5	48	85.7	864	3	DYN1_HUMAN	DYNAMIN-1.	4.53e+00
6	48	85.7	865	3	DYN1_CAEEL	DYNAMIN.	4.53e+00
7	48	85.7	883	3	DYN1_MOUSE	DYNAMIN (SHIBIRE PROT	4.53e+00
8	47	83.9	403	11	YNR5_YEAST	HYPOTHETICAL 45.7 KD	7.21e+00
9	47	83.9	642	5	HS72_PICAN	HEAT SHOCK PROTEIN 70	7.21e+00
10	47	83.9	644	5	HS71_HANPO	HEAT-SHOCK PROTEIN 70	7.21e+00
11	47	83.9	646	5	HS70_THEAN	HEAT SHOCK 70 KD PROT	7.21e+00
12	47	83.9	866	3	DYN2_MOUSE	DYNAMIN 2 (DYNAMIN UD	7.21e+00
13	47	83.9	870	3	DYN2_RAT	DYNAMIN 2.	7.21e+00
14	47	83.9	870	3	DYN2_HUMAN	DYNAMIN 2.	7.21e+00
15	46	82.1	181	10	VE4_HPV04	PROBABLE E4 PROTEIN.	1.14e+01
16	46	82.1	213	10	VE4_HPV65	PROBABLE E4 PROTEIN.	1.14e+01
17	45	80.4	201	10	Y007_NPVAC	HYPOTHETICAL 23.6 KD	1.79e+01
18	45	80.4	483	3	ENP1_YEAST	ENP1 PROTEIN.	1.79e+01
19	45	80.4	1088	7	PGDS_RAT	ALPHA PLATELET-DERIVE	1.79e+01
20	45	80.4	1089	7	PGDS_HUMAN	ALPHA PLATELET-DERIVE	1.79e+01
21	45	80.4	1089	7	PGDS_MOUSE	ALPHA PLATELET-DERIVE	1.79e+01
22	44	78.6	153	11	YJGD_HABIN	HYPOTHETICAL PROTEIN	2.79e+01

23	44	78.6	488	2	COAT_CAMVC	COAT PROTEIN.	2.79e+01
24	44	78.6	488	2	COAT_CAMVE	COAT PROTEIN.	2.79e+01
25	44	78.6	488	2	COAT_CAMVN	COAT PROTEIN.	2.79e+01
26	44	78.6	489	2	COAT_CAMVD	COAT PROTEIN.	2.79e+01
27	44	78.6	490	2	COAT_CAMVD	COAT PROTEIN.	2.79e+01
28	44	78.6	514	10	VS14_TRYBB	VARIANT SURFACE GLYCO	2.79e+01
29	44	78.6	640	5	HS71_ANOAL	HEAT SHOCK PROTEIN 70	2.79e+01
30	44	78.6	640	5	HS72_ANOAL	HEAT SHOCK PROTEIN 70	2.79e+01
31	44	78.6	762	9	SLAP_ACEKI	CELL SURFACE PROTEIN	2.79e+01
32	44	78.6	780	9	TREB_YEAST	PROBABLE TREHALASE (E	2.79e+01
33	44	78.6	929	3	DPOM_MAIZE	DNA POLYMERASE (EC 2.	2.79e+01
34	43	76.8	165	10	Y051_HABIN	HYPOTHETICAL PROTEIN	4.31e+01
35	43	76.8	497	10	VE2_HPV20	REGULATORY PROTEIN E2	4.31e+01
36	43	76.8	512	10	VP40_HSV7J	CARSID PROTEIN P40 (C	4.31e+01
37	43	76.8	793	3	DCMA_MEYSO	CARBON MONOXIDE DEHID	4.31e+01
38	43	76.8	1953	1	BN1L_YEAST	BN11 PROTEIN (SYNTHET	4.31e+01
39	42	75.0	205	10	YAB4_YEAST	HYPOTHETICAL 23.7 KD	6.61e+01
40	42	75.0	336	11	YDB1_SCHPO	HYPOTHETICAL 38.4 KD	6.61e+01
41	42	75.0	458	11	YGB4_YEAST	HYPOTHETICAL 54.6 KD	6.61e+01
42	42	75.0	593	11	YAG7_YEAST	HYPOTHETICAL 68.8 KD	6.61e+01
43	42	75.0	620	11	YJB9_YEAST	HYPOTHETICAL 72.4 KD	6.61e+01
44	42	75.0	1487	11	YGU7_YEAST	HYPOTHETICAL 167.1 KD	6.61e+01
45	42	75.0	2493	11	YBA4_YEAST	HYPOTHETICAL 287.5 KD	6.61e+01

ALIGNMENTS

RESULT	ID	MA29_DERFA	STANDARD;	PRT;	145 AA.
AC	P39674;				
DT	01-FEB-1995 (REL. 31, CREATED)				
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)				
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)				
DE	ALLERGEN MAG29 (FRAGMENT).				
GN	MAG29.				
OS	DERMATOPHAGOIDES FARINAE (HOUSE-DUST MITE).				
OC	EUKARYOTA; METAZOA; ARTHROPODA; CHELICERATA; ARACHNIDA; ACARI.				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-17.				
RX	MEDLINE; 94334283.				
RA	AKI T., FUJIKAWA A., WADA T., JYO T., SHIGETA S., MUROOKA Y., OKA S.,				
RA	ONO K.;				
RL	J. BIOCHEM. 115:435-440(1994).				
CC	-1- SIMILARITY: BELONGS TO HEAT SHOCK PROTEIN 70 FAMILY.				
DR	EMBL; D17676; G666007; -.				
DR	PIR; JX0313; JX0313.				
KW	ATP-BINDING; ALLERGEN.				
FT	NON_TER				
SQ	SEQUENCE 145 AA; 15594 MW; E9DC03F7 CRC32;				

Query Match 85.7%; Score 48; DB 6; Length 145;
Best Local Similarity 71.4%; Pred. No. 4.53e+00;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 14 ykeeddk 20
||:|
QY 2 YKDDDDK 8

RESULT	ID	DYN3_RAT	STANDARD;	PRT;	848 AA.
AC	Q08877;				
DT	01-OCT-1994 (REL. 30, CREATED)				
DT	01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)				
DE	DYNAMIN 3 (DYNAMIN, TESTICULAR) (T-DYNAMIN).				
GN	DNM3 OR DYN3.				
OS	RATTUS NORVEGICUS (RAT).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; RODENTIA.				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-TESTIS;				


```

Best Local Similarity 71.4%; Pred. No. 4.53e+00;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 553 ykddeek 559
QY 2 YKDDDK 8
||||:|

RESULT 4
ID DYN1_MOUSE STANDARD; PRT; 861 AA.
AC P39053;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE DYNAMIN-1 (DYNAMIN BREDDN19).
DE DNM1 OR DNM
GN DNM1 OR DNM
OS MUS MUSCULUS (MOUSE).
OS EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OS EUTHERIA; RODENTIA.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-NIH SWISS:
RC STIEF A., DER PUTTEN H.;
RA SUBMITTED (AXX-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RL -!- FUNCTION: MICROTUBULE-ASSOCIATED FORCE-PRODUCING PROTEIN INVOLVED
CC IN PRODUCING MICROTUBULE BUNDLES AND ABLE TO BIND AND HYDROLYZE
CC GTP. MOST PROBABLY INVOLVED IN VESICULAR TRAFFICKING PROCESSES, IN
CC PARTICULAR ENDOCYTOSIS.
CC -!- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.
CC -!- SIMILARITY: BELONGS TO THE DYNAMIN FAMILY.
CC -!- SIMILARITY: CONTAINS A PH DOMAIN.
EMBL; L31397; G487857; -
DR PROSITE; PS00410; DYNAMIN.
DR PROSITE; PS50003; PH DOMAIN.
DR MOTOR PROTEIN; GTP-BINDING; MICROTUBULES; MULTIGENE FAMILY;
KW ENDOCYTOSIS.
KW NP_BIND 38 45 GTP (BY SIMILARITY).
FT NP_BIND 136 140 GTP (BY SIMILARITY).
FT NP_BIND 205 208 GTP (BY SIMILARITY).
FT DOMAIN 515 621 PH.
FT SEQUENCE 861 AA; 97366 MW; 5A5E7E80 CRC32;

Query Match 85.7%; Score 48; DB 3; Length 861;
Best Local Similarity 71.4%; Pred. No. 4.53e+00;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 549 ykddeek 555
QY 2 YKDDDK 8
||||:|

RESULT 5
ID DYN1_HUMAN STANDARD; PRT; 864 AA.
AC Q05193;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE DYNAMIN-1.
DE DNM1 OR DNM.
GN DNM1 OR DNM.
OS HOMO SAPIENS (HUMAN).
OS EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OS EUTHERIA; PRIMATES.
[1]
RN SEQUENCE FROM N.A.
RC MEDLINE; 93328758.
RX VAN DER BLIEK A.M., REDELMEIER T.E., TISDALE E.J., MEYEROWITZ E.M.,
RA SCHMID S.L.;
RA J. CELL BIOL. 122:553-563(1993).
[2]
RN STRUCTURE BY NMR OF 511-630.
RX MEDLINE; 95153276.
RA DOWNING A.K., DRISCOLL P.C., GOUT I., SALIM K., ZVELEBIL M.J.,
RA WATERFIELD M.D.;

```

CURR. BIOL. 4:884-891(1994).
CC -!- FUNCTION: MICROTUBULE-ASSOCIATED FORCE-PRODUCING PROTEIN INVOLVED
CC IN PRODUCING MICROTUBULE BUNDLES AND ABLE TO BIND AND HYDROLYZE
CC GTP. MOST PROBABLY INVOLVED IN VESICULAR TRAFFICKING PROCESSES, IN
CC PARTICULAR ENDOCYTOSIS.
CC -!- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.
CC -!- SIMILARITY: BELONGS TO THE DYNAMIN FAMILY.
CC -!- SIMILARITY: CONTAINS A PH DOMAIN.
DR EMBL: L07807; G181849; -
DR EMBL: L07808; G181851; ALT_SEQ.
DR EMBL: L07809; G181853; ALT_SEQ.
DR EMBL: L07810; G181855; ALT_SEQ.
DR PROSITE: PS00410; DYNAMIN.
DR PROSITE: PS00003; PH DOMAIN.
KW MOTOR PROTEIN; GTP-BINDING; MICROTUBULES; MULTIGENE FAMILY;
FT ENDOCYTOSIS.
FT NP_BIND 38 45 GTP (BY SIMILARITY).
FT NP_BIND 136 140 GTP (BY SIMILARITY).
FT NP_BIND 205 208 GTP (BY SIMILARITY).
FT DOMAIN 519 625 PH.
SQ SEQUENCE 864 AA; 97407 MW; 97E7D339 CRC32;
Query Match 85.7%; Score 48; DB 3; Length 864;
Best Local Similarity 71.4%; Pred. No. 4.53e+00;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 553 ykdeek 559
Qy 2 YKDDDK 8
||||:|
RESULT 6
ID DYN1 CAEEL STANDARD; PRT; 865 AA.
AC P39055;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE DYNAMIN.
GN DYN-1.
OS CAENORHABDITIS ELEGANS.
OS EUKARYOTA; METAZOA; ACLOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA DER BLIEK A.M., SHURLAND D.L., MEYEROWITZ E.M.;
RL SUBMITTED (XXX-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -!- FUNCTION: MICROTUBULE-ASSOCIATED FORCE-PRODUCING PROTEIN INVOLVED
CC IN PRODUCING MICROTUBULE BUNDLES AND ABLE TO BIND AND HYDROLYZE
CC GTP. MOST PROBABLY INVOLVED IN VESICULAR TRAFFICKING PROCESSES, IN
CC PARTICULAR ENDOCYTOSIS.
CC -!- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.
CC -!- SIMILARITY: BELONGS TO THE DYNAMIN FAMILY.
CC -!- SIMILARITY: CONTAINS A PH DOMAIN.
DR EMBL: L29031; G456286; -
DR PROSITE: PS00410; DYNAMIN.
DR PROSITE: PS00003; PH DOMAIN.
KW MOTOR PROTEIN; GTP-BINDING; MICROTUBULES; MULTIGENE FAMILY;
FT ENDOCYTOSIS.
FT NP_BIND 40 47 GTP (BY SIMILARITY).
FT NP_BIND 138 142 GTP (BY SIMILARITY).
FT NP_BIND 207 210 GTP (BY SIMILARITY).
FT DOMAIN 519 624 PH.
SQ SEQUENCE 865 AA; 97156 MW; 3D916F54 CRC32;
Query Match 85.7%; Score 48; DB 3; Length 865;
Best Local Similarity 71.4%; Pred. No. 4.53e+00;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 552 ykdeek 558
Qy 2 YKDDDK 8
||||:|

RESULT 7
ID DYN DROME STANDARD; PRT; 883 AA.
AC P27619;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE DYNAMIN (SHIBIRE PROTEIN).
GN SHI.
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OS EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CANTON-S;
RX MEDLINE: 91260878.
RA CHEN M.S., OBAR R.A., SCHROEDER C.C., AUSTIN T.W., POODRY C.A.,
RA WADSWORTH S.C., VALLEE R.B.;
RL NATURE 351:583-586(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-OREGON-R;
RX MEDLINE: 91238973.
RA VAN DER BLIEK A.M., MEYEROWITZ E.M.;
RL NATURE 351:411-414(1991).
CC -!- FUNCTION: MICROTUBULE-ASSOCIATED FORCE-PRODUCING PROTEIN WHICH IS
CC INVOLVED IN THE PRODUCTION OF MICROTUBULE BUNDLES AND WHICH IS
CC ABLE TO BIND AND HYDROLYZE GTP. SHIBIRE IS IMPLICATED IN ENDOCYTIC
CC PROTEIN SORTING.
CC -!- DISEASE: SHIBIRE MUTATION IS THE CAUSE OF TEMPERATURE-SENSITIVE
CC PARALYSIS. THIS IS BELIEVED TO BE DUE TO A REVERSIBLE BLOCK OF
CC ENDOCYTOSIS, WHICH PREVENTS MEMBRANE CYCLING AND THUS DEPLETES
CC SYNAPTIC VESICLES.
CC -!- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.
CC -!- ALTERNATIVE PRODUCTS: AT LEAST THREE ISOFORMS ARE CREATED BY
CC ALTERNATIVE SPLICING OF THE SHI GENE.
CC -!- SIMILARITY: BELONGS TO THE DYNAMIN FAMILY.
CC -!- SIMILARITY: CONTAINS A PH DOMAIN.
DR EMBL: X59449; G7833; -
DR EMBL: X59448; G7831; -
DR EMBL: X59435; G7909; -
DR EMBL: X59436; G683556; -
DR PIR: S15413; S15413
DR PIR: S15497; S15497.
DR PIR: S15498; S15498.
DR PIR: S16130; S16130.
DR PIR: S17974; S17974.
DR PIR: S17975; S17975.
DR HSP: P19959; 1CLG.
DR FLYBASE: FEGN000392; SHI.
DR PROSITE: PS00410; DYNAMIN.
DR PROSITE: PS00003; PH DOMAIN.
KW MOTOR PROTEIN; GTP-BINDING; MICROTUBULES; ALTERNATIVE SPLICING;
KW ENDOCYTOSIS.
FT NP_BIND 33 40 GTP (POTENTIAL).
FT NP_BIND 131 135 GTP (POTENTIAL).
FT NP_BIND 200 203 GTP (POTENTIAL).
FT DOMAIN 513 621 PH.
FT DOMAIN 750 833 PRO-RICH.
FT VARSPLIC 635 640 MISSING (IN THIRD FORM).
FT VARSPLIC 836 836 V -> R (IN SHORT FORM).
FT VARSPLIC 837 883 MISSING (IN SHORT FORM).
FT VARIANT 141 141 G -> S (IN SHI-TS2 MUTANT).
FT VARIANT 268 268 G -> D (IN SHI-TS1 MUTANT).
FT VARIANT 594 594 R -> K.
SQ SEQUENCE 883 AA; 98537 MW; 489F2ED4 CRC32;
Query Match 85.7%; Score 48; DB 3; Length 883;
Best Local Similarity 71.4%; Pred. No. 4.53e+00;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 547 ykdeek 553
Qy 2 YKDDDK 8
||||:|

RESULT 8
 ID YN5 YEAST STANDARD; PRT: 403 AA.
 AC P53883;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 45.7 KD PROTEIN IN RPS3-PSD1 INTERGENIC REGION.
 UN YNL175C OR N1665.
 OS SACCCHAROMYCES CREVISIAE (BAKER'S YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA OBERMAIER B., PIRAVANDI E., RINKE M., DOWDEY H.;
 RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RNP).
 DR EMBL; Z71451; E239547;
 KW HYPOTHETICAL PROTEIN; NUCLEAR PROTEIN; RNA-BINDING.
 FT DOMAIN 144 149 RNA-BINDING (RNP2) (BY SIMILARITY).
 FT DOMAIN 183 190 RNA-BINDING (RNP1) (BY SIMILARITY).
 FT DOMAIN 241 246 RNA-BINDING (RNP2) (BY SIMILARITY).
 FT DOMAIN 280 287 RNA-BINDING (RNP1) (BY SIMILARITY).
 SQ SEQUENCE 403 AA; 45667 MW; D3D18998 CRC32;

Query Match 83.9%; Score 47; DB 11; Length 403;
 Best Local Similarity 50.0%; Pred. No. 7.21e+00;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 79 eykedaek 86
 ||:|:|
 QY 1 YKDDDDK 8

RESULT 9
 ID HS72 PICAN STANDARD; PRT: 642 AA.
 AC P53623; P53422;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE HEAT SHOCK PROTEIN 70 2.
 UN HSA2.
 OS PICHIA ANGUSTA (YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-ATCC 34438;
 RA DIESEL A., ROGGENKAMP R.;
 RL SUBMITTED (FEB-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
 CC -1- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO HEAT SHOCK PROTEIN 70 FAMILY.
 DR EMBL; U49932; G1256759;
 KW CHAPERONE; HEAT SHOCK; ATP-BINDING; MULTIGENE FAMILY; ACETYLATION.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
 SQ SEQUENCE 642 AA; 69941 MW; 997C782A CRC32;

Query Match 83.9%; Score 47; DB 5; Length 642;
 Best Local Similarity 57.1%; Pred. No. 7.21e+00;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 522 ykeedek 528
 ||:|:|
 QY 2 YKDDDDK 8

RESULT 10
 ID HS71 HANPO STANDARD; PRT: 644 AA.
 AC P53421;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE HEAT-SHOCK PROTEIN 70 1 (HSP72).
 GN HSA1.
 OS HANSENULA POLYMORPHA (YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN-ATCC 34438;
 RX MEDLINE; 96437974.
 RA TITORENKO V.I., EVERS M.E., DIESEL A., SAMYN B., VAN BEEMEN J.,
 RA ROGGENKAMP R.R., KIEL J.A.K.W., VAN DER KLEI I., VEENHUIS M.;
 RL YEAST 12:849-857(1996).
 CC -1- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
 CC -1- INDUCTION: BY HEAT SHOCK.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO HEAT SHOCK PROTEIN 70 FAMILY.
 DR EMBL; Z29379; G443915;
 KW CHAPERONE; HEAT SHOCK; ATP-BINDING; MULTIGENE FAMILY; ACETYLATION.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
 SQ SEQUENCE 644 AA; 70006 MW; D6C7BFA4 CRC32;

Query Match 83.9%; Score 47; DB 5; Length 644;
 Best Local Similarity 57.1%; Pred. No. 7.21e+00;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 524 ykeedek 530
 ||:|:|
 QY 2 YKDDDDK 8

RESULT 11
 ID HS70 THEAN STANDARD; PRT: 646 AA.
 AC P16019;
 DT 01-APR-1990 (REL. 14, CREATED)
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE HEAT SHOCK 70 KD PROTEIN (HSP 70.1).
 OS THEILERIA ANNULATA.
 OC EUKARYOTA; PROTOZOA; APICOMPLEXA; SPOKOZOA; COCCIDIA; PIROPLASMIDA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 90136713.
 RA MASON P.J., SHIELDS B.R., TAIT A., BECK P., HALL R.;
 RL MOL. BIOCHEM. PARASITOL. 37:27-36(1989).
 CC -1- SIMILARITY: BELONGS TO HEAT SHOCK PROTEIN 70 FAMILY.
 DR EMBL; J04653; G161869;
 DR HSP; P19120; INGG.
 DR PROSITE; PS00297; HSP70_1.
 DR PROSITE; PS00329; HSP70_2.
 DR PROSITE; PS01036; HSP70_3.
 KW ATP-BINDING; HEAT SHOCK.
 SQ SEQUENCE 646 AA; 70973 MW; C2EB974B CRC32;

Query Match 83.9%; Score 47; DB 5; Length 646;
 Best Local Similarity 57.1%; Pred. No. 7.21e+00;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 526 ykeedek 532
 ||:|:|
 QY 2 YKDDDDK 8

RESULT 12
 ID DYN2_MOUSE STANDARD; PRT: 866 AA.
 AC P39054;
 DT 01-FEB-1995 (REL. 31, CREATED)
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE DYNAMIN 2 (DYNAMIN UDNK).
 GN DNM2 OR DYN2.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.

KW SEQUENCE FROM N.A.
 RC STRAIN-NIH SWISS;
 RA STIEF A., DER PUTTEN H.;
 RL SUBMITTED (XXX-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: MICROTUBULE-ASSOCIATED FORCE-PRODUCING PROTEIN INVOLVED
 CC IN PRODUCING MICROTUBULE BUNDLES AND ABLE TO BIND AND HYDROLYZE
 CC GTP. MOST PROBABLY INVOLVED IN VESICULAR TRAFFICKING PROCESSES, IN
 CC PARTICULAR ENDOCYTOSIS.
 CC -1- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.
 CC -1- TISSUE SPECIFICITY: BELONGS TO THE DYNAMIN FAMILY.
 CC -1- SIMILARITY: CONTAINS A PH DOMAIN.
 DB EMBL: L31398; G487874; -
 DR PROSITE: PS00410; DYNAMIN.
 KW PROSITE: PS50003; PH_DOMAIN.
 KW MOTOR PROTEIN; GTP-BINDING; MICROTUBULES; MULTIGENE FAMILY;
 KW ENDOCYTOSIS.
 FT NP_BIND 38 45 GTP (BY SIMILARITY).
 FT NP_BIND 136 140 GTP (BY SIMILARITY).
 FT NP_BIND 205 208 GTP (BY SIMILARITY).
 FT DOMAIN 515 621 PH.
 SQ SEQUENCE 866 AA; 97683 MW; 5CB0DE3B CRC32;
 Query Match 83.9%; Score 47; DB 3; Length 866;
 Best Local Similarity 57.1%; Pred. No. 7.21e+00;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 Db 549 ykdeek 555
 |||:::
 QY 2 YKDDDDK 8
 RESULT 13
 ID DYN2_RAT STANDARD; PRT; 870 AA.
 AC P39052;
 DT 01-FEB-1995 (REL. 31, CREATED)
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE DYNAMIN 2.
 GN DNM2 OR DYN2.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY;
 RX MEDLINE: 94119943.
 COOK T.A., URRUTIA R., MCNIVEN M.A.;
 PROC. NATL. ACAD. SCI. U.S.A. 91:644-648(1994).
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE: 94140890.
 RA SONTAG J.-M., FYKSE E.M., USHKARYOV Y., LIU J.-P., ROBINSON P.J.,
 RA SUEHDHOF T.C.;
 RL J. BIOL. CHEM. 269:4547-4554(1994).
 CC -1- FUNCTION: MICROTUBULE-ASSOCIATED FORCE-PRODUCING PROTEIN INVOLVED
 CC IN PRODUCING MICROTUBULE BUNDLES AND ABLE TO BIND AND HYDROLYZE
 CC GTP. MOST PROBABLY INVOLVED IN VESICULAR TRAFFICKING PROCESSES, IN
 CC PARTICULAR ENDOCYTOSIS.
 CC -1- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.
 CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED, INCLUDING THE BRAIN.
 CC -1- HIGHEST LEVELS IN THE TESTIS.
 CC -1- ALTERNATIVE PRODUCTS: THREE FORMS OF DYN2 CAN BE PRODUCED BY
 CC ALTERNATIVE SPLICING. THE FORM SHOWN (IIBA) DIFFERS FROM I1AA
 CC BY THE REPLACEMENT OF AN EXON BY ANOTHER ONE WHICH IS HIGHLY
 CC SIMILAR.
 CC -1- SIMILARITY: BELONGS TO THE DYNAMIN FAMILY.
 CC -1- SIMILARITY: CONTAINS A PH DOMAIN.
 DR EMBL: L25605; G416396; -
 DR EMBL: L24562; G404073; -
 DR PROSITE: PS00410; DYNAMIN.
 DR PROSITE: PS50003; PH_DOMAIN.

KW MOTOR PROTEIN; GTP-BINDING; MICROTUBULES; MULTIGENE FAMILY;
 KW ENDOCYTOSIS; ALTERNATIVE SPLICING.
 FT NP_BIND 38 45 GTP (BY SIMILARITY).
 FT NP_BIND 136 140 GTP (BY SIMILARITY).
 FT NP_BIND 205 208 GTP (BY SIMILARITY).
 FT DOMAIN 519 625 PH.
 FT VARSPLIC 407 444
 LAFEAIKKVKKVKKPKLCKVDIVLQELISTVRQCTS ->
 MAFEAIKKVKKVKKPKLCKVDIVLQELISTVRQCTS
 (IN FORM I1AA).
 FT VARSPLIC 516 519
 MISSING (IN FORM IIC).
 FT CONFLICT 298 298 S -> T (IN REF. 1).
 FT CONFLICT 389 389 S -> T (IN REF. 1).
 FT CONFLICT 487 487 N -> K (IN REF. 1).
 FT CONFLICT 637 637 G -> E (IN REF. 1).
 FT CONFLICT 719 719 MISSING (IN REF. 1).
 FT CONFLICT 786 791 GPTPGP -> PHTGA (IN REF. 1).
 SQ SEQUENCE 870 AA; 98230 MW; 18544D0A CRC32;
 Query Match 83.9%; Score 47; DB 3; Length 870;
 Best Local Similarity 57.1%; Pred. No. 7.21e+00;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 Db 553 ykdeek 559
 |||:::
 QY 2 YKDDDDK 8
 RESULT 14
 ID DYN2_HUMAN STANDARD; PRT; 870 AA.
 AC P50570;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE DYNAMIN 2.
 GN DNM2 OR DYN2.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 96011652.
 RA DIATLOFF-ZITO C., GORDON A.J.E., DUCHAUD E., MERLIN G.;
 RL GENE 163:301-306(1995).
 CC -1- FUNCTION: MICROTUBULE-ASSOCIATED FORCE-PRODUCING PROTEIN INVOLVED
 CC IN PRODUCING MICROTUBULE BUNDLES AND ABLE TO BIND AND HYDROLYZE
 CC GTP. MOST PROBABLY INVOLVED IN VESICULAR TRAFFICKING PROCESSES, IN
 CC PARTICULAR ENDOCYTOSIS.
 CC -1- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.
 CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED.
 CC -1- SIMILARITY: BELONGS TO THE DYNAMIN FAMILY.
 CC -1- SIMILARITY: CONTAINS A PH DOMAIN.
 DR EMBL: L36983; G1196423; -
 KW MOTOR PROTEIN; GTP-BINDING; MICROTUBULES; MULTIGENE FAMILY;
 KW ENDOCYTOSIS; ALTERNATIVE SPLICING.
 FT NP_BIND 38 45 GTP (BY SIMILARITY).
 FT NP_BIND 136 140 GTP (BY SIMILARITY).
 FT NP_BIND 205 208 GTP (BY SIMILARITY).
 FT DOMAIN 519 625 PH.
 FT VARSPLIC 516 519 MISSING (IN A FORM).
 SQ SEQUENCE 870 AA; 98018 MW; 16C963CA CRC32;
 Query Match 83.9%; Score 47; DB 3; Length 870;
 Best Local Similarity 57.1%; Pred. No. 7.21e+00;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 Db 553 ykdeek 559
 |||:::
 QY 2 YKDDDDK 8
 RESULT 15
 ID VE4_HPV04 STANDARD; PRT; 181 AA.
 AC Q07852;

DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DE 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE PROBABLE E4 PROTEIN.
OS HUMAN PAPILLOMAVIRUS TYPE 4.
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; PAPOVAVIRIDAE; PAPILLOMAVIRUSES.

RP SEQUENCE FROM Nucleotide

RA MEDLINE; 93275568.

RA EGAWA K., DELIUS H., MATSUKURA T., KAWASHIMA M., DE VILLIERS E.M.;

RL VIROLOGY 194:789-799(1993).

DR EMBL; X70827; G312089; -.

KW EARLY PROTEIN.

ST DOMAIN 133 142 POLY-GLU

SQ SEQUENCE 181 AA; 20849 MW; 2769F837 CRC32;

Query Match

Best Local Similarity 82.1%; Score 46; DB 10; Length 181;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

112 dydedek 119

||:|:|

1 DYKDDDDK 8

Search completed: Tue Feb 3 15:32:01 1998
Job time : 13 secs.

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligenetics, Inc.

MParch_pp protein - protein database search, using Smith-Waterman algorithm

Run: Tue Feb 3 15:32:58 1998; MasPar time 2.41 Seconds
Tabular output not generated. 46.200 Million cell updates/sec

Title: >US-08-713-928A-10
Description: (1-8) from US08713928A.pep
Perfect Score: 56
Sequence: 1 DYKDDDDK 8

Scoring table: PAM 150
Gap 15

Searched: 111726 seqs, 13889129 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-genesq30
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23

Statistics: Mean 15.190; Variance 46.181; scale 0.329

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	56	100.0	8 23	W00971	FLAG epitope.	6.57e+00
2	56	100.0	8 22	W18236	FLAG epitope.	6.57e+00
3	56	100.0	8 15	R91066	Affinity tag peptide.	6.57e+00
4	56	100.0	8 15	R87022	Flag octapeptide.	6.57e+00
5	56	100.0	8 16	R91045	Antigenic N-terminal	6.57e+00
6	56	100.0	8 2	P70024	N-terminal fusion con	6.57e+00
7	56	100.0	8 3	R09327	Sequence of N-termina	6.57e+00
8	56	100.0	13 23	W00969	FLAG epitope with nuc	6.57e+00
9	56	100.0	13 3	P60023	Sequence of a Di-Palm	6.57e+00
10	56	100.0	27 9	R46933	Hyperglycosylated hil	6.57e+00
11	56	100.0	28 1	P82355	Fusion protein for ex	6.57e+00
12	56	100.0	75 23	W26361	Human pancreatic secr	6.57e+00
13	56	100.0	78 23	W26363	hPSTI-OMTKI3 chimeric	6.57e+00
14	56	100.0	78 23	W60034	hPSTI-SSImutant chime	6.57e+00
15	56	100.0	84 23	W00945	CMV500-4heptadCREB (N	6.57e+00
16	56	100.0	96 23	W26362	hPSTI-SSI chimeric su	6.57e+00
17	56	100.0	97 23	W00949	CMV500-4heptadFos leu	6.57e+00
18	56	100.0	105 23	W00947	CMV500-FosbZIP (Mo) pr	6.57e+00
19	56	100.0	113 23	W00950	CMV400-JunbZIP leucin	6.57e+00
20	56	100.0	122 16	R90842	Recombinant flag simi	6.57e+00

21	56	100.0	125 16	R78367	Human IL-3 mutant for	6.57e+00
22	56	100.0	125 16	R78366	Human IL-3 mutant for	6.57e+00
23	56	100.0	265 9	R52864	Glycophorin antibody	6.57e+00
24	56	100.0	273 9	R52865	Anti-Influenza N10 sc	6.57e+00
25	56	100.0	280 9	R52861	Ly-2+Ly-3 V domain w1	6.57e+00
26	56	100.0	284 7	R38321	Sequence of pSC49FLAG	6.57e+00
27	56	100.0	326 19	W06413	Flag tag/VT1 A subuni	6.57e+00
28	56	100.0	329 19	W06414	Flag tag/VT2 A subuni	6.57e+00
29	56	100.0	396 22	W18574	Aggrecanase artificia	6.57e+00
30	56	100.0	396 22	W18575	Aggrecanase artificia	6.57e+00
31	56	100.0	463 19	W05137	TGF alpha-ETA fusion	6.57e+00
32	56	100.0	496 16	R95058	GAL4-DT-IL-2 multido	6.57e+00
33	56	100.0	546 22	W18237	Human glucocerebrosid	6.57e+00
34	56	100.0	637 5	R26982	(FRP5)-ETA fusion pro	6.57e+00
35	56	100.0	637 5	R26983	(FRP51)-ETA fusion pr	6.57e+00
36	56	100.0	651 19	W05135	scFv(225)-ETA fusion	6.57e+00
37	56	100.0	651 19	W05136	scFv(FRP5)-ETA fusion	6.57e+00
38	56	100.0	699 19	W05138	scFv(FRP5)/TGF alpha	6.57e+00
39	56	100.0	832 4	R22603	Taq polymerase encode	6.57e+00
40	56	100.0	832 4	R22604	Taq polymerase encode	6.57e+00
41	56	100.0	892 19	W05140	scFv2(225)/FRP5)-ETA	6.57e+00
42	56	100.0	892 19	W05143	scFv2(FRP5)/FRP5)-ETA	6.57e+00
43	56	100.0	895 19	W05142	scFv2(FRP5)/FRP5)-ETA	6.57e+00
44	56	100.0	899 19	W05144	scFv2(FRP5)/FRP5)-ETA	6.57e+00
45	56	100.0	1020 19	W05141	scFv2(FRP5/225)-ETA (6.57e+00

ALIGNMENTS

RESULT 1
ID W00971 standard; Peptide; 8 AA.
AC W00971;
DT 12-NOV-1997 (first entry)
DE FLAG epitope.
KW DNA binding protein; RNA binding protein; amphipathic peptide;
KW acidic extension peptide; gene control; gene regulation;
KW transcription; dominant negative protein; cancer; drug therapy;
KW drug design; leucine zipper; FLAG.
OS Synthetic.
PN W09705249-A2.
PD 13-FEB-1997.
PF 31-JUL-1996; U12590.
PR 31-JUL-1996; US-001654.
PR 31-JUL-1996; US-001654.
PR 29-MAY-1996; US-018496.
PA (KRYL/) KRYLOV D.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (VINS/) VINSON C R.
PI Krylov D, Vinson CR;
DR WPI; 97-145687/13.
DR New nucleic acid binding proteins - having an acidic amino acid
PT sequence extension at the amino-terminus, to increase ability to
PT regulate gene transcription, useful e.g. in cancer therapeutics
PS Claim 16, Page 58; 144pp; English.
CC This peptide sequence comprises a FLAG epitope that can be
CC attached to the N-terminus of a nucleic acid binding protein
CC (NABP) such as a leucine zipper bZIP protein or bHLH protein.
CC Claimed NABPs having an appended acidic extension peptide (see
CC W00958-65) can regulate the function of a target nucleic acid or
CC gene to which they are bound, and act as potent dominant-negative
CC regulators of gene transcription, cell growth and cell proliferation.
CC They can be used in cancer therapeutics, to treat diseases caused by
CC eukaryotic microorganisms or by viruses, and as tools for drug
CC development, rational drug design, and drug and gene therapies.
CC Sequence 8 AA;
SQ

Query Match 100.0%; Score 56; DB 23; Length 8;

Best Local Similarity 100.0%; Pred. No. 6.57e+00;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 dykdddk 8

Qy 1 DYKDDDDK 8

2
 ID W18236 standard; Peptide; 8 AA.
 AC W18236;
 DT 01-OCT-1997 (first entry)
 DE FLAG epitope.
 KW Lysosomal enzyme; lysosome; transgenic plant; glucocerebrosidase;
 KW alpha-L-iduronidase; IUDA; enzyme replacement therapy;
 KW Gaucher disease; Hurler syndrome; FLAG.
 OS Synthetic.
 PN WO9710353-A1.
 PD 20-MAR-1997.
 PF 13-SEP-1996; U14730.
 PR 14-SEP-1995; US-003737.
 PA (CROP-) CROPTECH DEV CORP.
 PA (VIRG) VIRGINIA TECH INTELLECTUAL PTY INC.
 PI Cramer CL, Oishi KK, Radin DN, Weissborn DL;
 DR WPI: 97-202248/18.
 PT Production of enzymatically active (modified) lysosomal enzyme in
 transgenic plants - useful in treatment of lysosomal storage
 disorders
 CC Claim 6; Page 37; English.
 CC The FLAG epitope (W18236) is utilised in novel constructs for
 CC expression of lysosomal enzymes, e.g. human glucocerebrosidase
 CC and alpha-L-iduronidase in transgenic plants. The epitope coding
 CC sequence is fused in-frame to the C-terminus of the lysosomal
 CC enzyme coding sequence (see also W17153) in order to facilitate the
 CC detection and purification of the gene product (see also W18237).
 CC It is designed to be a hydrophilic marker peptide situated on a
 CC protein surface to facilitate antibody interactions.
 CC Sequence 8 AA;
 SQ
 Query Match 100.0%; Score 56; DB 22; Length 8;
 Best Local Similarity 100.0%; Pred. No. 6.57e+00;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 dykdddk 8
 QY 1 DYKDDDK 8
 3
 ID R91066 standard; Peptide; 8 AA.
 AC R91066;
 DT 23-MAY-1996 (first entry)
 DE Affinity tag peptide.
 KW Interleukin-1 type-3 receptor; IL-1-3R; Immune-associated disease;
 KW vector; antibody; therapy; affinity tag.
 OS Synthetic.
 PN WO9607739-A2.
 PD 14-MAR-1996.
 PF 11-SEP-1995; U12037.
 PR 09-SEP-1994; US-303957.
 PA (NEUR-) NEUROCRINE BIOSCIENCES INC.
 PI Cleverger W, DeSouza EB, Liaw CW, Lovenberg TW;
 PI Oltersdorf T;
 DR WPI: 96-171614/17.
 PT Interleukin-1 type 3 receptor proteins - useful for the treatment of
 PT immune-associated diseases
 PS Disclosure; Page 7; 64pp; English.
 CC An affinity tag (R91066) may be linked to interleukin-1 type-3
 CC receptors (see e.g. R91064 and R91054) produced by recombinant
 CC DNA technology. The peptide facilitates purification of the
 CC expressed recombinant protein.
 CC Sequence 8 AA;
 SQ
 Query Match 100.0%; Score 56; DB 15; Length 8;
 Best Local Similarity 100.0%; Pred. No. 6.57e+00;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 dykdddk 8
 QY 1 DYKDDDK 8

QY 1 DYKDDDK 8
 4
 ID R87022 standard; peptide; 8 AA.
 AC R87022;
 DT 11-JUN-1996 (first entry)
 DE Flag octapeptide.
 KW B2LF2; Epstein-Barr virus; EBV; C-type lectin; beta chain; MHC; antigen;
 KW major histocompatibility complex; immunoglobulin; cytotoxic T cell;
 KW autoimmune disease; myasthenia gravis; multiple sclerosis; allergy;
 KW systemic lupus erythematosus; organ transplant rejection; asthma; IL-7;
 KW tissue transplant rejection; therapy; cancer; viral disease; mouse;
 KW Interleukin-7.
 OS Synthetic.
 PN WO9530015-A2.
 PD 09-NOV-1995.
 PF 28-APR-1995; U05348.
 PR 28-APR-1994; US-235397.
 PA (IMMV) IMMUNEX CORP.
 PA (UMOR) UNIV MISSOURI.
 PA (USSH) US NAT INST OF HEALTH.
 PI Alderson M, Armitage RJ, Cohen JL, Comeau MR, Farrah TM;
 PI Hutt-fletcher LM, Spriggs MK;
 DR WPI: 95-393086/50.
 PT Epstein-Barr virus B2LF2 fusion proteins - used for treating e.g.
 PT auto-immune disease, transplant rejection, allergy, asthma, cancer
 PT or viral infection.
 PS Example 1; Page 38; 51pp; English.
 CC This sequence represents the Flag octapeptide, and was used in creating a
 CC B2LF2-immunoglobulin Fc fusion protein (B2LF2/Fc). B2LF2 is a
 CC mouse interleukin-7 (IL-7) leader sequence (see R87021), an
 CC immunoglobulin Fc region (see R87023) and a flexible linker (see R87024)
 CC are joined to the extracellular domain (residues 34 to 223) of the B2LF2
 CC family. The C-type lectin domain is found in type II membrane proteins.
 CC The B2LF2 protein is capable of binding the beta chain of a major
 CC histocompatibility complex (MHC) class II antigen. Fusion proteins with
 CC an oligomerising zipper domain (OZD), instead of an immunoglobulin Fc
 CC region, can also be created. B2LF2 proteins inhibit antigen-specific
 CC antibody formation, proliferation of blood mononuclear cells and
 CC cytotoxic T cell responses. They also exhibit superantigen-like
 CC activity. The proteins can be used for treating or preventing autoimmune
 CC diseases such as myasthenia gravis, multiple sclerosis and systemic lupus
 CC erythematosus. Also, for treating organ or tissue transplant rejection
 CC and for treating or preventing allergy or asthma. They can be used for
 CC treating cancer and viral disease, especially EBV infection.
 CC Sequence 8 AA;
 SQ
 Query Match 100.0%; Score 56; DB 15; Length 8;
 Best Local Similarity 100.0%; Pred. No. 6.57e+00;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 dykdddk 8
 QY 1 DYKDDDK 8
 5
 ID R91045 standard; Peptide; 8 AA.
 AC R91045;
 DT 01-AUG-1996 (first entry)
 DE Antigenic N-terminal peptide for fusion to shuIL-1R.
 KW Interleukin-1 receptor; human; soluble; N-terminal peptide; antigen;
 KW epitope; shuIL-1R; monoclonal antibody; bovine mucosal enterokinase;
 KW interleukin-1; IL-1; immune response; mammal; diagnosis; therapy;
 KW regulation; immune disease; inflammatory disease.
 OS Synthetic.
 PN U5542888-A.
 PD 20-FEB-1996.
 PF 25-NOV-1987; 125627.
 PR 25-NOV-1987; US-125627.

PR 25-FEB-1988; US-160550.
 PR 13-OCT-1988; US-258756.
 PR 21-DEC-1989; US-453488.
 PR 17-JUN-1992; US-904071.
 PA (IMMUNEX) IMMUNEX CORP.
 PI Dower SK, March CJ, Sims JE, Urdal DL;
 DR WPI: 96-150236/15.
 PT Use of sol. IL-1 receptors to suppress IL-1-mediated immune
 PT responses - e.g. for treatment of inflammation in mammals, esp.
 PT humans
 PS Disclosure; Column 8; 22pp; English.
 CC This sequence represents an antigenic N-terminal peptide epitope. This
 CC sequence can be fused to soluble human interleukin-1 receptor (shuIL-1R).
 CC By using this sequence, recombinant shuIL-1R can be assayed and purified
 CC easily, by using a monoclonal antibody that reversibly binds this
 CC epitope. By using bovine mucosal enterokinase, this sequence can be
 CC cleaved from the recombinant shuIL-1R. Proteins capped with this peptide
 CC may also be resistant to intracellular digestion in E. coli. The
 CC shuIL-1R can be used in a method for suppressing interleukin-1 (IL-1)
 CC mediated immune responses in a mammal. The method comprises
 CC administering an effective amount of shuIL-1R to the mammal, preferably
 CC in a dose of 500ng-5mg/kg/day. The shuIL-1R is effective for use in
 CC assay, diagnosis or therapy for regulation of immune or inflammatory
 CC activities, in contrast to membrane-bound full-length mature IL-1
 CC receptors.
 SQ Sequence 8 AA;

Query Match 100.0%; Score 56; DB 16; Length 8;
 Best Local Similarity 100.0%; Pred. No. 6.57e+00;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 dykdddk 8
 |||||
 QY 1 DYKDDDDK 8

RESULT 6
 ID P70024 standard; peptide; 8 AA.
 AC P70024;
 DT 12-FEB-1991 (first entry)
 DE N-terminal fusion construct comprising mutant human G-CSF and
 DE yeast alpha-factor leader sequence.
 KW Granulocyte-colony stimulating factor; fusion protein;
 KW yeast alpha-factor leader sequence.
 PN EP-243153-A.
 PD 28-OCT-1987.
 PF 22-APR-1987; 303509.
 PI 22-APR-1986; US-856643.
 PI 14-NOV-1986; US-931458.
 PI IMMUNEX CORP.
 PI Cosman DJ, Gillis S, Mochizuki DY, March CJ, Price VL,
 PI Tushinski RJ, Urdal DL.
 DR WPI: 87-300791/43.
 PR New-pure-human-granulocyte colony stimulating factor and muteins -
 PR obt'd. by recombinant DNA methods in high yields, used for potentiating
 PR immune responses, treating leukaemias etc.
 PS Disclosure; Page 13; 37pp; English.
 CC The sequence encodes a fusion construct which links a mutant
 CC sequence of human granulocyte-colony stimulating factor to
 CC a yeast alpha-factor leader sequence. Fusion proteins capped
 CC with this peptide are resistant to intracellular degradation.
 CC See also N70029, N70031-35 and P70025.
 SQ Sequence 8 AA;

Query Match 100.0%; Score 56; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 6.57e+00;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 dykdddk 8
 |||||
 QY 1 DYKDDDDK 8

RESULT 7
 ID R09327 standard; peptide; 8 AA.
 AC R09327;
 DT 29-MAR-1992 (first entry)
 DE Sequence of N-terminal octapeptide for compsn. contg. human
 DE interleukin-3 (IL-3) analogue.
 KW Lymphokine; bone marrow proliferation; cytopenia therapy.
 PN W09001039-A.
 PD 08-FEB-1990.
 PF 14-JUN-1989; U02599.
 PR 20-JUL-1988; US-221699.
 PA (IMMU-) IMMUNEX CORP.
 PI Anderson DM, Cosman DJ, Price VL;
 DR WPI: 90-067162/09.
 PT Compsn. contg. recombinant non-glycosylated human interleukin-3
 PT - has increased biological activity and binding affinity, for
 PT treating cytopenias
 PS Claim 5; Page 18; 23pp; English.
 CC The inventors claim a pharmaceutical compsn. which contains an
 CC effective amt. of a recombinant human interleukin-3 protein analogue,
 CC rhIL-3, (Asp15, Asp70). The rhIL-3 analogue has AA SQ in R09326.
 CC The compsn. may also comprise the N-terminal octapeptide in R09327,
 CC and a diluent and 1 or more than 1 biological response modifier.
 CC The compsn. has a biological specific activity of equal to or more
 CC than 4.0 x 10 to the 7 mcg/mg in a human bone marrow proliferation
 CC assay, and a binding affinity for human monocyte IL-3 receptors of
 CC equal to or more than 4.0 x 10 to the 10 (M to the minus 1).
 SQ Sequence 8 AA;

Query Match 100.0%; Score 56; DB 3; Length 8;
 Best Local Similarity 100.0%; Pred. No. 6.57e+00;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 dykdddk 8
 |||||
 QY 1 DYKDDDDK 8

RESULT 8
 ID W00969 standard; Peptide; 13 AA.
 AC W00969;
 DT 12-NOV-1997 (first entry)
 DE FLAG epitope with nuclear localisation sequence.
 DE DNA binding protein; RNA binding protein; amphipathic peptide;
 KW acidic extension peptide; gene control; gene regulation;
 KW transcription; dominant negative protein; cancer; drug therapy;
 KW drug design; EBP; leucine zipper; FLAG.
 OS Synthetic.
 PN W09705249-A2.
 PD 13-FEB-1997.
 PF 31-JUL-1996; U12590.
 PR 31-JUL-1996; US-001654.
 PR 31-JUL-1995; US-001654.
 PR 29-MAY-1996; US-018496.
 PA (KRYL/) KRYLOV D.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (VINS/) VINSON C R.
 PI Krylov D, Vinson CR;
 DR WPI: 97-145687/13.
 PR New nucleic acid binding proteins - having an acidic amino acid
 PR sequence extension at the amino-terminus, to increase ability to
 PR regulate gene transcription, useful e.g. in cancer therapeutics
 PS Claim 16; Page 54; 144pp; English.
 CC This peptide sequence comprises a FLAG epitope with a nuclear
 CC localisation sequence. It is used in novel modified nucleic
 CC acid binding protein (NABPs) chimeras. These claimed NABPs, esp.
 CC leucine zipper proteins and BHLH proteins, have an appended acidic
 CC extension peptide (see W00958-65). They can regulate the function
 CC of a target nucleic acid or gene to which they are bound, and act
 CC as potent dominant-negative regulators of gene transcription, cell
 CC growth and cell proliferation. They can be used in cancer
 CC therapeutics, to treat diseases caused by eukaryotic microorganisms
 CC or by viruses, and as tools for drug development, rational drug

CC ✓ design, and drug and gene therapies.
SQ Sequence 13 AA;

Query Match 100.0%; Score 56; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 6.57e+00;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 dykdddk 9
|||||
Qy 1 DYKDDDDK 8

RESULT 9
ID P60023 standard; peptide; 13 AA.
AC P60023; 1991 (first entry)
DE Sequence of a Di-Palmityl Derivatized Identification Peptide.
KW Hybrid protein; purification.
FH Key Location/Qualifiers
FT Misc_difference 11
FT /label= palmityl-Lys
FT /label= palmityl-Lys
FT /label= palmityl-Lys

PN EP-195680-A.
PD 24-SEP-1986.
PF 20-MAR-1986; 302079.
PR 21-MAR-1985; US-714691.
PA (IMMU-) IMMUNEX CORP.
PI Hopp TP, Conlon PJ, Bektesh SL, March CJ;
DR WPI; 86-253627/39.
PT New identification peptide contg. antigenic and cleavage portions
PT - useful for making antibodies to purify recombinant hybrid
PT protein by affinity chromatography
PS Example; Page 13; 23pp; English.
CC The identification peptides of the invention comprise: (1) an
CC N-terminal sequence corresp. to the antigen portion of a natural
CC protein and/or synthetic protein to which a monoclonal antibody has
CC been raised; and (2) a second portion which can be cleaved at a
CC specified AA residue by a particular proteolytic agent. Pref.
CC component (2) can be linked to a protein (1) which is to be purified.
CC In IP-1, the sequence Asp-Tyr-Lys constitutes the antigenic portion
CC of the peptide while the sequence Asp-Asp-Asp-Lys constitutes the
CC protease cleavable linking portion.
SQ Sequence 13 AA;

Query Match 100.0%; Score 56; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 6.57e+00;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 dykdddk 8
|||||
Qy 1 DYKDDDDK 8

RESULT 10
ID R46933 standard; Protein; 27 AA.
AC R46933;
DE Hyperglycosylated hIL-2 N-terminal fragment.
KW Human; Interleukin-2; IL-2; N-linked carbohydrate; biotinylation;
KW Alpha-factor; signal peptide; 2micron plasmid; detection.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..6
FT /note= "Alpha-factor leader peptide"
FT Peptide 7..15
FT /note= "identification peptide"
FT Modified_site 18..21
FT /note= "N-linked glycosylation site"
FT Modified_site 22..25
FT /note= "N-linked glycosylation site"
FT Peptide 26..27
FT /note= "hIL-2 N-terminal fragment"

PN US5298395-A.
PD 29-MAR-1994.
PF 25-APR-1989; 343471.
PR 25-APR-1989; US-343471.
PR 24-JAN-1992; US-827517.
PR 05-NOV-1992; US-972010.
PA (IMMV) IMMUNEX CORP.
PI Park LS;
DR WPI; 94-100334/12.
DR N-PSDB; Q46933.
PT Cell detection method for cells expressing cytokine receptor -
PT using hyperglycosylated recombinant cytokine conjugated to a
PT detectable functional moiety via oligosaccharide residue
PS Example 3; Column 13-14; 8pp; English.
CC This sequence is encoded by part of a plasmid which was used in the
CC production of a human interleukin-2 (IL-2) protein which has
CC increased levels of N-linked carbohydrate for biotinylation. This
CC sequence spans the Asp718 site at amino acid 79 near the
CC 3' end of the alpha-factor signal peptide to the SpeI site in the
CC 2micron plasmid sequence. This polypeptide may be used in
CC the detection method of the invention for the diagnosis of particular
CC physiological or pathological conditions.
SQ Sequence 27 AA;

Query Match 100.0%; Score 56; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.57e+00;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 dykdddk 14
|||||
Qy 1 DYKDDDDK 8

RESULT 11
ID P82355 standard; protein; 28 AA.
AC P82355;
DT 13-NOV-1990 (first entry)
DE Fusion protein for expression of hIL-3.
KW Interleukin-3; fusion protein; yeast; alpha-factor leader; KEX2;
KW flag peptide; cytopenia; haematopoiesis; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT Region 1..6
FT /label= alpha-factor leader
FT /note= "C-terminal 5 AAs from D718"
FT Peptide 7..14
FT /label= identification peptide
FT /note= "N-terminal antigenic flag"
FT Cleavage_site 14..15
FT /note= "cleaved by bovine mucosal enterokinase"
FT CDS 15..28
FT /label= hIL-3 N-terminal
PN W08805469-A.
PD 28-JUL-1988.
PF 11-JAN-1988; U00011.
PR 20-JAN-1987; US-004466.
PA (IMMU-) Immunex Corp.
PI Anderson DM, Cosman DJ, Price VL;
DR WPI; 88-220313/31.
DR N-PSDB; N82115.
PT Purified recombinant human interleukin-3 - used to potentiate immune
PT response or assist in reconstituting normal blood following
PT haematopoietic cell suppression.
PS Disclosure; 2pp; English.
CC The sequence is encoded by 4 oligomers. The 8 residue "flag"-DYKDDDDK-
CC is highly antigenic and provides an epitope reversibly bound by MAb,
CC enabling rapid assay and easy purific. of the expressed recombinant hIL-3
CC The sequence is also cleaved by bovine mucosal enterokinase after the
CC second Lys. Fusion proteins capped with this peptide are also resistant
CC to intracellular degradation prior to secretion. DNA encoding fragment
CC can be used to construct an expression plasmid, pEC125, which can then be
CC used to transform a yeast expression strain, XV2181, which is cultured
CC to produce high levels of recombinant hIL-3. The protein can be used

CC to treat various cytopenias, or can be used in compsns. to potentiate
CC immune response or assist in reconstituting blood cell populations
CC following viral infection or radiation- or chemotherapy-induced
CC haematopoietic cell suppression.
CC See also P82334.
SQ Sequence 28 AA;

Query_Match 100.0%; Score 56; DB 1; Length 28;
Best Local Similarity 100.0%; Pred. No. 6.57e+00;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 dykdddk 14
| | | | |
QY 1 DYKDDDK 8

RESULT 12
ID W26361 standard; Protein; 75 AA.

AC W26361;
DT 02-DEC-1997 (first entry)
DE Human pancreatic secretory trypsin inhibitor (M2 tagged).
KW Subtilisin inhibitor; protease inhibitor; PSTI; hPSTI.M2; human;
KW pancreatic secretory trypsin inhibitor; enzyme engineering;
KW protein engineering; baculovirus; detergent.
OS Chimeric Homo sapiens;
OS Chimeric synthetic.

FH Key Location/Qualifiers
FT Protein 1..65
FT /label= hPSTI
FT Peptide 66..75
FT /label= M2_tag
FT /note= "facilitates cloning and affinity
FT purification"

PN W09715670-A1.
PD 01-MAY-1997.

PF 25-OCT-1996; U17153.
PR 25-OCT-1995; US-548186.
PA (ARRI-) ARRIS PHARM CORP.
PI Estell DA, Fisher J, Hartman C, McGrath M, Schmidt B;
DR WPI: 97-259023/23.
DR N-PSDB; T84523.

PT Modified subtilisin inhibitor - having altered reactive site
PT resulting in decreased immunogenicity, used in detergent
PT composition(s)

PS Example 1; Fig 8; 69pp; English.
CC This sequence comprises human pancreatic secretory trypsin
CC inhibitor (PSTI), modified to include a C-terminal M2 affinity tag.
CC It is encoded by a DNA construct (see T84523) suitable for
CC baculoviral expression. Novel modified subtilisin inhibitors are
CC claimed in which a human standard mechanism inhibitor, such as
CC PSTI, has at least one amino acid (aa) of the reactive site
CC substituted with different aa so that its dissociation constant for
CC inhibition, for at least one subtilisin, is reduced by at least a
CC factor of 100. Also claimed are: a nucleic acid (1) encoding the
CC subtilisin inhibitor; expression vectors containing (1); and host
CC cells containing (1). The subtilisin inhibitors (see W26362-64)
CC combine the low allergenicity of human standard mechanism inhibitors,
CC which are not specific for subtilisin, and the high affinity of
CC non-human subtilisin inhibitors. They form a complex with
CC subtilisin, so as to avoid problems of allergenicity in detergent
CC formulations, and can also be used in affinity purification and
CC (diagnostic) quantification of subtilisin.
SQ Sequence 75 AA;

Query Match 100.0%; Score 56; DB 23; Length 75;
Best Local Similarity 100.0%; Pred. No. 6.57e+00;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 68 dykdddk 75
| | | | |
QY 1 DYKDDDK 8

RESULT 13
ID W26363 standard; Protein; 78 AA.
AC W26363;
DT 02-DEC-1997 (first entry)
DE hPSTI.OMTKY3 chimeric subtilisin inhibitor.
KW Subtilisin inhibitor; protease inhibitor; human; PSTI;
KW hPSTI.OMTKY3; pancreatic secretory trypsin inhibitor;
KW turkey ovomucoid third domain protein; enzyme engineering;
KW protein engineering; detergent.
OS Chimeric Homo sapiens;
OS Chimeric Meleagris gallopavo;
OS Chimeric synthetic.
FH Key Location/Qualifiers
FT Peptide 1..3
FT /label= Sig_peptide
FT Protein 4..78
FT /label= Mat_protein
FT Peptide 22..31
FT /label= OMTKY3
FT /note= "turkey ovomucoid third domain protein
FT reactive site"
FT Peptide 66..78
FT /label= M2_tag
FT /note= "affinity tag facilitates purification"
PN W09715670-A1.
PD 01-MAY-1997.
PF 25-OCT-1996; U17153.
PR 25-OCT-1995; US-548186.
PA (ARRI-) ARRIS PHARM CORP.
PI Estell DA, Fisher J, Hartman C, McGrath M, Schmidt B;
DR WPI: 97-259023/23.
DR N-PSDB; T84525.
PT Modified subtilisin inhibitor - having altered reactive site
PT resulting in decreased immunogenicity, used in detergent
PT composition(s)
PS Example 2; Fig 10A; 69pp; English.
CC This polypeptide comprises human pancreatic secretory trypsin
CC inhibitor (PSTI), modified to include a turkey ovomucoid third
CC domain (OMTKY3) reactive site. It is encoded by a DNA construct
CC (T84525) produced by PCR amplification of overlapping synthetic
CC oligonucleotides. Novel modified subtilisin inhibitors are
CC claimed in which a human standard mechanism inhibitor, such as
CC PSTI, has an altered reactive site so that its dissociation
CC constant for inhibition of subtilisin is reduced by at least a
CC factor of 100. Also claimed are: a nucleic acid encoding the
CC subtilisin inhibitor; expression vectors; and transformed host
CC cells. The novel subtilisin inhibitors combine the low
CC allergenicity of human standard mechanism inhibitors, which are
CC not specific for subtilisin, and the high affinity of non-human
CC subtilisin inhibitors such as OMTKY3. They form a complex with
CC subtilisin, so as to avoid problems of allergenicity in detergent
CC formulations, and can also be used in affinity purification and
CC (diagnostic) quantification of subtilisin.
SQ Sequence 78 AA;

Query Match 100.0%; Score 56; DB 23; Length 78;
Best Local Similarity 100.0%; Pred. No. 6.57e+00;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 71 dykdddk 78
| | | | |
QY 1 DYKDDDK 8

RESULT 14
ID W26364 standard; Protein; 78 AA.
AC W26364;
DT 02-DEC-1997 (first entry)
DE hPSTI.SS mutant chimeric subtilisin inhibitor.
KW Subtilisin inhibitor; protease inhibitor; human; PSTI.SSI;
KW pancreatic secretory trypsin inhibitor; human; PSTI.SSI;
KW protein engineering; detergent.
OS Chimeric Homo sapiens;

OS Chimeric Streptomyces sp.;
 FH Chimeric synthetic.
 FT Key Location/Qualifiers
 FT Peptide 1..3
 FT /label= Sig_peptide
 FT Protein 4..78
 FT /label= Mat_protein
 FT Peptide 22..32
 FT /label= SSI
 FT /note= "Streptomyces subtilisin inhibitor mutated
 FT reactive site contains Ala at P6, Ala at
 FT P5, Ala at P4, Leu at P1 and Arg at P3."
 FT Peptide 67..78
 FT /label= M2 tag
 FT /note= "affinity tag facilitates purification"
 FT WO9715670-A1.
 PD 01-MAY-1997.
 PD 25-OCT-1996; U17153.
 PD 25-OCT-1995; US-548186.
 PA (ARRI-) ARRIS PHARM CORP.
 PA Estell DA, Fisher J, Hartman C, McGrath M, Schmidt B;
 PI WPI; 97-259023/23.
 DR R-PSDB; T84526.
 PT Modified subtilisin inhibitor - having altered reactive site
 PT resulting in decreased immunogenicity, used in detergent
 PT composition(s)
 PS Example 2; Fig 11A; 69pp; English.
 CC This polypeptide comprises human pancreatic secretory trypsin
 CC inhibitor (PSI), modified to include a Streptomyces subtilisin
 CC inhibitor (SSI) mutated reactive site. It is encoded by a DNA
 CC construct (T84526) produced by PCR amplification of overlapping
 CC synthetic oligonucleotides. Novel modified subtilisin inhibitors
 CC are claimed in which a human standard mechanism inhibitor, such as
 CC PSI, has an altered reactive site so that its dissociation
 CC constant for inhibition of subtilisin is reduced by at least a
 CC factor of 100. Also claimed are: a nucleic acid encoding the
 CC subtilisin inhibitor; expression vectors; and transformed host
 CC cells. The novel subtilisin inhibitors combine the low
 CC allergenicity of human standard mechanism inhibitors, which are
 CC not specific for subtilisin, and the high affinity of non-human
 CC subtilisin inhibitors such as SSI. They form a complex with
 CC subtilisin, so as to avoid problems of allergenicity in detergent
 CC formulations, and can also be used in affinity purification and
 CC (diagnostic) quantification of subtilisin.
 SQ Sequence 78 AA;

Query Match 100.0%; Score 56; DB 23; Length 78;
 Best Local Similarity 100.0%; Pred. No. 6.57e+00;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 71 dykdddk 78
 |||||
 QY 1 DYKDDDK 8

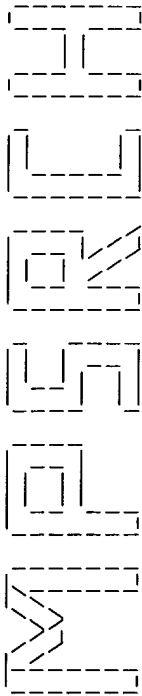
RESULT 15
 ID W00945 standard; Protein; 84 AA.
 DT 11-NOV-1997 (first entry)
 DE CMV500-4heptadCREB (New4hepCREB).
 KW DNA binding protein; RNA binding protein; amphipathic peptide;
 KW acidic extension peptide; gene control; gene regulation;
 KW transcription; dominant negative protein; CREB; cancer;
 KW drug therapy; drug design; CMV.
 OS Chimeric Homo sapiens;
 OS Chimeric cytomegalovirus;
 OS Chimeric synthetic.
 FH Key Location/Qualifiers
 FT /note= "encoded by GAC"
 FT WO9705249-A2.
 PD 13-FEB-1997.
 PF 31-JUL-1996; U12590.

PR 31-JUL-1996; US-001654.
 PR 31-JUL-1995; US-001654.
 PR 29-MAY-1996; US-018496.
 PA (KRYL/) KRYLOV D.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (VINS/) VINSON C R.
 PI Krylov D, Vinson CR;
 DR WPI; 97-145687/13.
 DR N-PSDB; T84345.
 PT New nucleic acid binding proteins - having an acidic amino acid
 PT sequence extension at the amino-terminus, to increase ability to
 PT regulate gene transcription, useful e.g. in cancer therapeutics
 PS Claim 16; Page 83-84; 144pp; English.
 CC This sequence comprises the polypeptide CMV500-4heptadCREB, or
 CC New4hepCREB, which has a 4heptad appended acidic extension.
 CC Claimed nucleic acid binding proteins (NABPs) such as CREB that
 CC have acidic peptide extensions are capable of regulating the
 CC function of a target nucleic acid or gene to which they are bound,
 CC and act as potent dominant-negative regulators of gene
 CC transcription, cell growth and cell proliferation. They can be
 CC used in cancer therapeutics, to treat diseases caused by eukaryotic
 CC microorganisms or by viruses, and as tools for drug development,
 CC rational drug design, and drug and gene therapies. They have an
 CC extended protein interaction surface or multimerisation or
 CC dimerisation interface that increases the stability of complexes
 CC formed.
 SQ Sequence 84 AA;

Query Match 100.0%; Score 56; DB 23; Length 84;
 Best Local Similarity 100.0%; Pred. No. 6.57e+00;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 dykdddk 9
 |||||
 QY 1 DYKDDDK 8

Search completed: Tue Feb 3 15:33:16 1998
 Job time : 18 secs.



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

on: Tue Feb 3 20:06:16 1998; MasPar time 146.36 Seconds
Tabular output not generated. 892.149 Million cell updates/sec

Title: >US-08-713-928A-5
Description: (1-463) from US08713928A.seq
Perfect Score: 463
N.A. Sequence: 1 CAAATACGATATACCGAATA.....CCGGTGAAGTAAGCAGSTC 463
Comp: GTATGCTATAATAGGCTTAT.....GCCCACTTCATTCGTCWAG

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 397346 seqs, 141010104 bases x 2

Post-processing: Minimum Match 08
Listing first 45 summaries

Database: EST-C
1:EST197 2:EST198 3:EST199 4:EST200 5:EST201 6:EST202
7:EST203 8:EST204 9:EST205 10:EST206 11:EST207 12:EST208
13:EST209 14:EST210 15:EST211 16:EST212 17:EST213
18:EST214 19:EST215 20:EST216 21:EST217 22:EST218
23:EST219 24:EST220 25:EST221 26:EST222 27:EST223
28:EST224 29:EST225 30:EST226 31:EST227 32:EST228
33:EST229 34:EST230 35:EST231 36:EST232 37:EST233
38:EST234 39:EST235 40:EST236 41:EST237 42:EST238
43:EST239 44:EST240 45:EST241 46:EST242 47:EST243
48:EST244 49:EST245 50:EST246 51:EST247 52:EST248
53:EST249 54:EST250 55:EST251 56:EST252 57:EST253
58:EST254 59:EST255 60:EST256 61:EST257 62:EST258
63:EST259 64:EST260 65:EST261 66:EST262 67:EST263
68:EST264 69:EST265 70:EST266 71:EST267 72:EST268
73:EST269 74:EST270 75:EST271 76:EST272 77:EST273
78:EST274 79:EST275 80:EST276 81:EST277 82:EST278
83:EST279 84:EST280 85:EST281 86:EST282 87:EST283
88:EST284 89:EST285 90:EST286 91:EST287 92:EST288
93:EST289 94:EST290 95:EST291 96:EST292 97:EST293
98:EST294
EST-D
99:EST295 100:EST296 101:EST297 102:EST298 103:EST299
104:EST300 105:EST301 106:EST302 107:EST303 108:EST304
109:EST305 110:EST306 111:EST307 112:EST308 113:EST309
114:EST310 115:EST311 116:EST312 117:EST313 118:EST314
119:EST315

Database:

Statistics: Mean 10.704; Variance 4.122; scale 2.597

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	28	6.0	362	104	HUM039A01B	Human fetal brain cDN
2	28	6.0	506	91	C07172	Human fetal brain cDN
3	27	5.8	535	97	AA214564	zr92e05.sl Soares NH
4	26	5.6	231	34	AA254732	vr92e04.sl Soares mou
5	26	5.6	243	67	AA414924	vc50g04.rl Knowles So
6	26	5.6	327	13	AA243996	nc05a05.rl NCI CGAP P
7	26	5.6	342	47	AA271298	vb73a07.rl Soares mou
8	26	5.6	408	23	AA251149	zs04e04.sl Soares NH
9	26	5.6	513	68	C23818	Dictyostelium discoid
10	26	5.6	746	64	AA320277	LD11555.9prime LD Dro
11	25	5.4	127	16	AA098579	mw83f05.rl Stratagene
12	25	5.4	184	8	AA231547	mw31e07.rl Soares mou
13	25	5.4	247	68	C23711	Dictyostelium discoid
14	25	5.4	247	45	AA266722	Dictyostelium discoid
15	25	5.4	253	27	AA303340	EST15962 Aorta endoch
16	25	5.4	281	63	AA388781	vb25a08.rl Soares mou
17	25	5.4	288	38	AA330297	EST34001 Embryo, 12 w
18	25	5.4	292	1	AA178264	mt09a05.rl Soares mou
19	25	5.4	295	106	HUM354G10B	Human aorta cDNA 5'-e
20	25	5.4	335	45	AA266824	mc95e08.rl Soares mou
21	25	5.4	336	83	AA420493	nc60b12.sl NCI CGAP P
22	25	5.4	343	6	AA220439	mw68e01.rl Soares mou
23	25	5.4	352	15	AA250411	mw16g12.rl Soares mou
24	25	5.4	354	47	AA273430	vb97h12.rl Soares mou
25	25	5.4	366	84	AA424191	zr81c07.rl Soares tot
26	25	5.4	372	45	AA266978	mp99b03.rl Soares mou
27	25	5.4	379	18	AA116482	mp94e09.rl Soares 2Nb
28	25	5.4	427	48	AA275763	vc14f09.rl Barstead M
29	25	5.4	469	69	W88869	zh74f02.rl Soares fet
30	25	5.4	500	97	AA279725	z992d09.rl Soares NH
31	25	5.4	509	41	C22944	Dictyostelium discoid
32	25	5.4	841	117	C22769	Dictyostelium discoid
33	25	5.4	1159	68	C23669	Dictyostelium discoid
34	24	5.2	149	15	AA253130	zr52f01.rl Soares NH
35	24	5.2	174	45	AA265356	mo82d11.rl Bedington
36	24	5.2	180	102	CELK091F3R	C.elegans cDNA clone
37	24	5.2	285	47	AA272758	va42a05.rl Soares mou
38	24	5.2	334	40	AA337483	EST42469 Endometrial
39	24	5.2	377	68	C23739	Dictyostelium discoid
40	24	5.2	398	100	AA291844	zt45u04.rl Soares ova
41	24	5.2	400	46	AA267358	mr91d09.rl Soares mou
42	24	5.2	433	98	AA281722	zt06h03.sl Soares NH
43	24	5.2	579	7	AA244692	CpEST.133 psKilminusc
44	24	5.2	633	116	AA457791	vf72a10.rl Soares mou
45	24	5.2	812	41	C22924	Dictyostelium discoid

ALIGNMENTS

RESULT	LOCUS	HUM039A01B	362 bp	mRNA	EST	21-MAY-1996
1	DEFINITION	Human fetal brain cDNA 5'-end	GEN-039A01.			
	ACCESSION	D51453				
	NID	9951689				
	KEYWORDS	EST(expressed sequence tag); Human fetal brain; similar to known (May 29, 1995).				
	SOURCE	Homo sapiens	CDNA to mRNA, clone lib:clontech human fetal brain			
	ORGANISM	Polya+ mRNA (#6535).				
		Homo sapiens				
		Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;				
		Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae;				
		Homo.				
	REFERENCE	1 (bases 1 to 362)				
	AUTHORS	Fujiwara, T., Hirano, H., Katagiri, T., Kawai, A., Kuga, Y., Nagata, M., Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y., Shinomiya, H., Takaichi, A., Takeda, S., Watanabe, T., Takahashi, E., Hirai, Y., Maekawa, H., Shin, S. and Nakamura, Y.				
	TITLE	Unpublished				

JOURNAL
 REFERENCE
 2 (bases 1 to 362)
 FUJIWARA,T.
 Direct Submission
 Submitted (30-MAY-1995) to the DDBJ/EMBL/GenBank databases. Tsutomu
 Fujiwara, Otsuka Gen Research Institute, Otsuka Pharmaceutical
 Co.,Ltd; 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01,
 Japan (Tel:0886-65-2888, Fax:0886-37-1035)
 Submitted (30-MAY-1995) to DDBJ by:
 Tsutomu Fujiwara
 Otsuka Gen Research Institute
 Otsuka Pharmaceutical Co.,Ltd
 463-10 kagasuno Kawauchi-cho
 Tokushima, Tokushima
 771-01
 Japan
 Phone: 0886-65-2888
 Fax : 0886-37-1035.

```

max . 0000 57 1035.
FEATURES
Location/Qualifiers
source
1..362

```

```

/organism="Homo sapiens"
/clone_lib="Clontech human fetal brain polyA+ mRNA
/#5535"

```

BASE COUNT		ORIGIN	
131 a	88 c	56 g	77 t
		10 others	

Query Match 6.0%; Score 28; DB 104; Length 362;
Best Local Similarity 62.5%; Pred. No. 4.53e-04;
Matches 45; Conservative 5; Mismatches 22; Indels 0; Gaps 0;

Db 268 aaaaaaaaaaaaaavvargaaaaaasataataaaaaa 327

[illegible]

QY 216 AAAAAATTAAAAAAGGAAGTATATATATGTAAGAATAACTCCATTCAAAATAT 215

Db 328 aaaaaaaaaa 339

RESULT	2	EST	08-AUG-1996
LOCUS	C07172	506 bp	mrna
DEFINITION	similar to none.		
ACCESSION	C07172		
NID	g1503948		
KEYWORDS	EST(expressed sequence tag).		
SOURCE	Rattus norvegicus rat pancreatic islets cdna to mrna, clone_lib:Rat		
	pancreatic islet cdna library:RBC865.		

ORGANISM
Rattus norvegicus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
Murinae; Rattus.
pancreatic islet cell carcinoma

REFERENCE	1 (sites)
1	1

AUTHORS Takeda, J.
TITLE Large scale collection of expressed sequence tags (ESTs) from rat pancreatic islet cDNA library

" JOURNAL Unpublished (1996)
REFERENCE 2 (bases 1 to 506)

AUTHORS
Takeda, J.
Direct Submissions

JOURNAL TITLE
Submitted (07-JUN-1996) to the DDBJ/EMBL/GenBank databases. Jun
Direct submission

Takeda, Institute for Molecular and cellular Regulation, Gunma

University, Dept. of Molecular Medicine; 3-39-15 Showa-machi,
Maebashi 371 Japan (E-mail: itakeda@new.sh.gunma-u.ac.jp)

maebashi 371, Japan (E-mail: jtakeeda@new.sov.yamada u.ac.jp,
Tel:81-272-20-8856, Fax:81-272-20-8896)

COMMENT	Project='Rat pancreatic islet cDNA'

vector=Lambda ZAP II
primer-T7 primer

Rritel-Eco RI

Rs site 2=Xho I
mpNA was prepared from normal rat islets. cDNA was directionally

mRNA was prepared from normal rat islets. cDNA was unidirectionally synthesized from the Xho I in the vector to the EcoRI site.

FEATURES	Location/Qualifiers
1. <i>Location</i>	500
2. <i>Qualifiers</i>	500

source 1.506

CONFIDENTIAL

1

```

/organism="Rattus norvegicus"
/clone="RBC865"
/clone_lib="Rat pancreatic islet cDNA"
/tissue_type="rat pancreatic islets"

BASE COUNT      185 a      68 C      63 g      190 t
ORIGIN

Query Match      6.0%; Score 28; DB 91; Length 506;
Best Local Similarity 70.6%; Pred. No. 4.53e-04;
Matches 48; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Db 180 ttcaaatTTTaatatgatataataattgattgtttgattacataattataattataa 239
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 70 TTCAAAATTTTAAATTTAATATCTACTTTCAACATATTATTATCCTAATTATCAAAATGC 129
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 240 tttaataga 247
      | ||||
QY 130 ATGTATGA 137

RESULT 3
LOCUS AA214564 535 bp mRNA EST 19-MAY-1997
DEFINITION z192e05.s1 Soares NbHTGBC Homo sapiens cDNA clone 683168 3'.
ACCESSION AA214564
NID G1813189
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrion eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.

REFERENCE 1 (bases 1 to 535)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevasakis,E.,
Waterson,R., Williamson,A., Wohlmann,P. and Wilson,R.
WashU-Merck EST Project
TITLE Unpublished (1995)

```

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1129 Std Error: 0.00
Seq primer: -41ml13 fwd. ET from Amersham
High quality sequence stop: 275.

```

FEATURES
source
Location/Qualifiers
1..535
/organism="Homo sapiens"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand of
was prepared from germinal B-cells (flow-sorted from
tonsils) provided by Dr. Louis Staudt of the NCI, and was
then primed with a Not I - oligo(dN) primer [5'
TGTTCACCATCTCAAGTGGGACGCCCTCATTTTTTTTTTTTTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/clone="683168"
/clone_lib="Soares NbHTGBC"
/tissue_type="Germinal B-cell"
/lab_host="DH10B"
complement(<1..535)
/db_xref="GBD:5586488"
153 a 123 c 101 g 151 t 7 others
BASE COUNT
ORIGIN

```

[illegible]

KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.

REFERENCE
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubouque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

TITLE The WashU-HMI Mouse EST Project
COMMENT Unpublished (1996)

Contact: Marria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:331481
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 110.

FEATURES

Location/Qualifiers	1..127
/organism="Mus musculus"	
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. M30 CD4+ cells. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGG 3' -3' adaptor sequence: 5' CTCAGTTTTTTTTTTT 3'"	
/clone="350689"	
/clone_lib="Stratagene mouse Tcell 937311"	
/dev_stage="M30 CD4+ cells"	
/lab_host="SOLR (kanamycin resistant)"	
<1..>127	13 a 4 c 5 g 105 t

mRNA

BASE COUNT	13 a	4 c	5 g	105 t
ORIGIN				

Query Match 5.4%; Score 25; DB 16; Length 127;
Best Local Similarity 64.7%; Pred. No. 5,63e-02;
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

```

Db      40  ttctttttttttttttttttttttttttttttttttttttttttttttttttttttttttt 99
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
287  TTTTTCATTATTTATTTGAATCGAGGTATTCTTTTACAATAATACCTTCTCTTTT 228

Db      100  ttttttttttttaaaaaaacattt 124
         ||||  ||||||  ||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Cp      227  TTTTAATTTTTTGAAATAAAGTTTT 203

```

RESULT 12

LOCUS	DEFINITION	Accession	NID	KEYWORDS	SOURCE	ORGANISM	EST
RA231547	mw31607.r1 Soares mouse 3NME12 5 Mus musculus cdNA clone 672324 5'	RA231547	g1853854	house mouse.	Mus musculus	Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	26-FEB-1997
REFERENCE							
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubouque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.							

[illegible]

TITLE Sexual cDNA in CSM
JOURNAL Unpublished (1997)
COMMENT PROJECT = "Dictyostelium discoideum CDNA project in Japan".
FEATURES Location/Qualifiers

source 1..224
 /organism="Dictyostelium discoideum"
 /strain="KAX3"
 /cell_type="Gamete"
 /clone_lib="EC"
 /dev_stage="Sexually competent"
BASE COUNT 149 a 11 c 11 g 53 t
ORIGIN

Query Match 5.4%; Score 25; DB 68; Length 224;
- Best Local Similarity 63.4%; Pred. NO. 5.63e-02;
Matches 59; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Db 99 gactattataaatgaatgtaaaaaaaaaaaaaaa 191
																								-			
																								-			
																								-			
																								-			
195 GCATATAAAGCTTTACTTCAAAAATTAAAAAAAGAAGATATATTGTAAGG 254																											
																								-			
																								-			
Bb 159 aaaaaaaaaaaaaaaaaaaaaaaaaaaa 191																											
QY 255 ATAACTCCATTCACAAATATATAATGAAGAAA 287																											
																								-			
																								-			

RESULT 14
LOCUS AA266722 247 bp mRNA EST 21-MAR-1997
DEFINITION mz98f06.r1 Soares mouse lymph node NbMLN Mus musculus CDNA clone
721475 5'.
ACCESSION AA266722
IID gl903564
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.

REFERENCE 1 (bases 1 to 247)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubouque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theisinger,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)

TITLE Contact: Marra M/Mouse EST Project
JOURNAL WashU-HMI Mouse EST Project
COMMENT Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314-286-9810
Email: mousee@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:446971
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 206.
Location/Qualifiers
source 1..247
 /organism="Mus musculus"
 /strain="C57BL/6j"

FEATURES /note=Vector: pTZ19-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; [5'
TGTTACAATCGTAGGGAGGCAGCATCTTTTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTZ19 vector. RNA provided by Dr. Bertrand Jordan. Library constructed and normalized by Bento Soares and M.Fátima Bonaldo."
/clone_lib="721475"

```

      mRNA      /dev_stage="adult"
BASE COUNT    91 a      44 c      59 g      58 t      1 others
ORIGIN

Query Match      5.4%; Score 25; DB 27; Length 253;
Best Local Similarity 83.8%; Pred. No. 5.63e-02;
Matches 31; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 171 attaaaaaaagagagatcatcagtgtaaaagata 207
      |||||
QY 221 ATTAAAAAAGAAAGTATTATTGTAAGAGATA 257
      |||||

Search completed: Tue Feb 3 20:14:24 1998
Job time : 488 secs.
```

WIREH

(TM)

Release 2.1D John F. Collins, BioComputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

on: Tue Feb 3 20:14:46 1998; MasPar time 150.66 Seconds
775.329 Million cell updates/sec
abular output not generated.

Title: >US-08-713-928A-5
Description: (1-463) from US08713928A.seq
Perfect Score: 463
N.A. Sequence: 1 CAATACGATATACCGAATA.....CCGGTAAAGTAAGCAGSTC 463
Comp: GTTATGCTATAATGGCTTAT.....GCCCACTTTCATTCGTCWAG

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 333433 seqs, 126143548 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

STS
1:STS1 2:STS2 3:STS3 4:STS4 5:STS5 6:STS6 7:STS7 8:STS8
9:STS9 10:STS10 11:STS11 12:STS12 13:STS13 14:STS14
15:STS15 16:STS16 17:STS17 18:STS18 19:STS19
20:STS20 21:STS21 22:STS22 23:STS23 24:STS24 25:STS25
26:STS26 27:STS27 28:STS28 29:STS29 30:STS30 31:STS31
32:STS32 33:STS33 34:STS34 35:STS35 36:STS36 37:STS37 38:STS38 39:STS39
40:STS40 41:STS41 42:STS42 43:STS43 44:STS44 45:STS45
46:STS46 47:STS47 48:STS48 49:STS49 50:STS50
51:STS51 52:STS52 53:STS53 54:STS54 55:STS55
56:STS56 57:STS57 58:STS58 59:STS59 60:STS60
61:STS61 62:STS62 63:STS63 64:STS64 65:STS65
66:STS66 67:STS67 68:STS68 69:STS69 70:STS70
71:STS71 72:STS72 73:STS73 74:STS74 75:STS75
76:STS76 77:STS77 78:STS78 79:STS79 80:STS80
81:STS81 82:STS82 83:STS83 84:STS84 85:STS85
86:STS86 87:STS87 88:STS88 89:STS89 90:STS90
91:STS91 92:STS92 93:STS93 94:STS94 95:STS95

Statistics: Mean 10.998; Variance 4.755; scale 2.313

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description	Pred. No.
C 1	29	6.3	488 94	PFAF10513 Plasmodium falciparum	1.32e-03
C 2	29	6.3	488 45	AF010513 Plasmodium falciparum	1.32e-03

3	28	6.0	506 46	DD1325844 Dictyostelium discoid	5.75e-03
4	28	6.0	607 41	AA5220111 Dictyostelium discoid	5.75e-03
5	28	6.0	607 88	MM1313659 Dictyostelium discoid	5.75e-03
6	27	5.8	245 11	HUMSWS29333 human chromosome 7 ST	2.44e-02
7	27	5.8	245 12	HUMSWS29333 human chromosome 7 ST	2.44e-02
8	27	5.8	400 4	G11046 human STS SHGC-9031 c	2.44e-02
9	27	5.8	441 86	MM1303648 Dictyostelium discoid	2.44e-02
10	27	5.8	441 35	AA510004 Dictyostelium discoid	2.44e-02
11	27	5.8	535 76	HSAA37572 Dictyostelium discoid	2.44e-02
12	27	5.8	535 14	AA214564 Dictyostelium discoid	2.44e-02
13	26	5.6	154 67	HS1284157 Dictyostelium discoid	1.00e-01
14	26	5.6	154 32	AA436647 Dictyostelium discoid	1.00e-01
15	26	5.6	200 80	HSAA7745 Dictyostelium discoid	1.00e-01
16	26	5.6	200 40	AA507745 Dictyostelium discoid	1.00e-01
17	26	5.6	306 46	DDC362 Dictyostelium discoid	1.00e-01
18	26	5.6	306 44	C24362 Dictyostelium discoid	1.00e-01
19	26	5.6	327 77	HSAA43996 Dictyostelium discoid	1.00e-01
20	26	5.6	327 25	AA243996 Dictyostelium discoid	1.00e-01
21	26	5.6	408 19	AA251149 Dictyostelium discoid	1.00e-01
22	26	5.6	408 50	HS1165877 Dictyostelium discoid	1.00e-01
23	26	5.6	632 46	CPAA32271 Dictyostelium discoid	1.00e-01
24	26	5.6	683 88	MM1313909 Dictyostelium discoid	1.00e-01
25	26	5.6	746 48	DMAA92077 Dictyostelium discoid	1.00e-01
26	26	5.6	746 48	DMAA92077 Dictyostelium discoid	1.00e-01
27	25	5.4	229 32	AA495388 Dictyostelium discoid	3.99e-01
28	25	5.4	229 48	DMAA95388 Dictyostelium discoid	3.99e-01
29	25	5.4	281 86	MM1302983 Dictyostelium discoid	3.99e-01
30	25	5.4	284 79	HSAA70491 Dictyostelium discoid	3.99e-01
31	25	5.4	294 41	AA521334 Dictyostelium discoid	3.99e-01
32	25	5.4	294 71	HS1312971 Dictyostelium discoid	3.99e-01
33	25	5.4	298 26	AA478272 Dictyostelium discoid	3.99e-01
34	25	5.4	298 64	HS1270075 Dictyostelium discoid	3.99e-01
35	25	5.4	336 26	AA420493 Dictyostelium discoid	3.99e-01
36	25	5.4	336 56	HS1215344 Dictyostelium discoid	3.99e-01
37	25	5.4	343 96	MM1139538 Dictyostelium discoid	3.99e-01
38	25	5.4	366 56	HS1223451 Dictyostelium discoid	3.99e-01
39	25	5.4	470 73	HS1322012 Dictyostelium discoid	3.99e-01
40	25	5.4	500 51	HS1184115 Dictyostelium discoid	3.99e-01
41	25	5.4	500 21	AA279725 Dictyostelium discoid	3.99e-01
42	25	5.4	718 44	C24618 Dictyostelium discoid	3.99e-01
43	25	5.4	718 46	DDC618 Dictyostelium discoid	3.99e-01
44	25	5.4			
45	25	5.4			

ALIGNMENTS

RESULT 1
ID PFAF10513 standard; DNA; STS; 488 BP.
AC AF010513;
NI 92253505
DT 14-JUL-1997 (Rel. 52, Created)
DT 14-JUL-1997 (Rel. 52, Last updated, Version 1)
DE Plasmodium falciparum microsatellite TA28 sequence.
KW STS.
OS Plasmodium falciparum (malaria parasite)
OC Eukaryotae; mitochondrial eukaryotes; Alveolata; Apicomplexa;
OC Haemosporidia; Plasmodium.
RN [1]
RP 1-488
RA Su X.Z., Wellem's T.E.;
RT "Toward a high-resolution Plasmodium falciparum linkage map:
RT polymorphic markers from hundreds of simple sequence repeats";
RL Genomics 33:430-444(1996).
RN [2]
RP 1-488
RA Su X.Z., Wellem's T.E.;
RT ;
RL Submitted (27-JUN-1997) to the EMBL/GenBank/DBJ databases.
RL NIAID-Laboratory of Parasitic Diseases, NIH, Bethesda, MD 20892,
RL USA
RL Key
FH Location/Qualifiers

```

RL (E-mail:gxurushi@bank.dna.afrc.go.jp, Tel:0298-53-4664,
RN Fax:0298-53-6614)
RN [2]
RA Saitoh T., Morio T., Tanaka Y., Ochiai H.;
RT "Developmental cDNA in Dictyostelium discoideum(970723)";
RT Unpublished.
CC PROJECT = "Dictyostelium discoideum cDNA project in Japan"
FH Key Location/Qualifiers
FH source
FH 1..506
FH /clone_lib="ss"
FH /dev_stage="Slug"
FH /organism="Dictyostelium discoideum"
FH /sequenced_mol="cDNA to mRNA"
FH /strain="AX4"
SQ Sequence 506 BP; 238 A; 63 C; 63 G; 142 T; 0 other;

Query Match 6.0%; Score 28; DB 46; Length 506;
Best Local Similarity 67.5%; Pred. No. 5.75e-03;
Matches 54; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Db 393 ttatattagaagaaaaataaaaaataaaaaataaaaaataaaaaataattataatt 452
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 208 TTCTACTCAAAATAATTAATAAAAAAGAAAGTATATTGTTAAAGATAATCTCCATT 267
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 453 taaataaaaaaa 472
||||| ||| ||| ||| |||

QY 268 CAAATATAAATGAAAAA 287

RESULT 4
LOCUS AA522011 607 bp mRNA EST 17-JUL-1997
DEFINITION vh78c09 r1 Knowles Solter mouse E6 5d whole embryo Mus musculus
cDNA clone 893104 3'.
ACCESSION AA522011
NID 92262756
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 607)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Thelings,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:521064
High quality sequence stop: 356.
Location/Qualifiers
1..607
/organism="Mus musculus"
/strain="B6D2 F1/J"
/note="Vector: pBluescribe (modified); Site_1: MluI;
Site_2: SalI; cloned unidirectionally from mRNA prepared
from 13,500 2-cell stage embryos Primer: SalI(dT):
5'-CGGTCGACCTGTCAGCGTTTTTTTTTTT-3'. cDNAs were
cloned into the MluI/SalI sites of a modified pBluescribe
vector using commercial linkers (NEB).
/clone="893104"
FEATURES
source

```

```
/clone_lib="Knowles Solter mouse E6 5d whole embryo"
/dev_stage="embryo (post-implantation)"
/lab_host="DH10B"
complement(<1..>607)
BASE COUNT      266 a 96 c 108 g 137 t
ORIGIN

Query Match
Best Local Similarity 66.0%; Score 28; DB 41; Length 607;
Matches 68; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

Db 330 agtttagtggaaaaataaaacatttcttctaaataaaataaaataaaataaaataaa 389
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 186 AGTCAACTTGACTATATAAAC-TTTACTTCAAAAATTAATAAAAGAGAGTATATT 244
.
Db 390 aaaaaataaaataaaataaaataaaataaaataaaataaaataaaataaa 432
| ||||| ||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 245 ATTGTAAGAGATAATACTCCATTCAAAATATAAAATGAAGAAAAA 287

JULT 5
MM1313659 standard; RNA; EST; 607 BP.
AC AA522011;
NI 92262756
DT 18-JUL-1997 (Rel. 52, Created)
DE vh8c09.r1 Knowles Solter mouse E6 5d whole embryo Mus musculus
DE cdna clone 893104 3'.
KW EST.
OS Mus musculus (house mouse)
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
OC Murinae; Mus.
RN [1]
RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA Waterston R.;
RT "The WashU-HMI Mouse EST Project";
RL Unpublished.
CC Contact: Marra M/Mouse EST Project WashU-HMI Mouse EST Project
CC Washington University School of MedicineP 4444 Forest Park Parkway,
CC Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC Email: mouseest@wustl.edu This clone is available
CC royalty-free through LNL; contact the IMAGE Consortium
CC (info@image.llnl.gov) for further information. MGI:521064 High
quality sequence stop: 356.
Key Location/Qualifiers
source 1..607
organism="Mus musculus"
strain="B6D2 Fl/J"
note="Vector: pBluescribe (modified); Site_1: MluI;
Site_2: SalI; Cloned unidirectionally from mRNA prepared
from 13,500 2-cell stage embryos. Primer: SalI(dT):
5'-CGGTCCACGTCGACGCTTTT-3', cDNAs were cloned
into the MluI/SalI sites of a modified pBluescribe vector
using commercial linkers (NEB)."
clone="893104"
clone_lib="Knowles Solter mouse E6 5d whole embryo"
dev_stage="embryo (post-implantation)"
lab_host="DH10B"
complement(<1..>607)
Sequence 607 BP; 266 A; 96 C; 108 G; 137 T; 0 other;
mrna
Query Match
Best Local Similarity 66.0%; Score 28; DB 88; Length 607;
Matches 68; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

Db 330 agtttagtggaaaaataaaacatttcttctaaataaaataaaataaaataaa 389
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 186 AGTCAACTTGACTATATAAAC-TTTACTTCAAAAATTAATAAAAGAGAGTATATT 244
Db 390 aaaaaataaaataaaataaaataaaataaaataaaataaaataaaataaa 432
| ||||| ||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 245 ATTGTAAGAGATAATACTCCATTCAAAATATAAAATGAAGAAAAA 287

RESULT 6
LOCUS HUMSWS2933 245 bp DNA STS 02-APR-1997
DEFINITION human chromosome 7 STS SWSS2933.
ACCESSION G16157
NID g1185334
KEYWORDS STS sequence; primer; sequence tagged site.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE 1 (bases 1 to 245)
AUTHORS Bouffard,G.G., Iyer,L.M., Idol,J.R., Braden,V.V., Cunningham,A.F.,
Weintraub,L.A., Mohr-Tidwell,R.M., Peluso,D.C., Fulton,R.S.,
Leckie,M.P. and Green,E.D.
TITLE A collection of 1814 human chromosome 7-specific STSs
MEDLINE 97189344
REFERENCE 2 (bases 1 to 245)
AUTHORS Green,E.D.
TITLE Human chromosome 7 STSs
JOURNAL Unpublished (1997)
COMMENT
GDB: GDB:1318428
Contact: Eric D. Green
Genome Technology Branch
National Human Genome Research Institute/NIH
49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
Tel: 3014020201
Fax: 3014024735
Email: egreen@nhgri.nih.gov
Primer A: TCAACATCTCTCAACAGAC
Primer B: AGTGTGAATGTTTGATGG
STS size: 162
PCR Profile:
Presoak: 0 degrees C for 0.00 minute(s)
Denaturation: 92 degrees C for 0.17 minute(s)
Annealing: 55 degrees C for 1.00 minute(s)
Polymerization: 72 degrees C for 1.00 minute(s)
PCR Cycles: 35
Thermal Cycler: PerkinElmer 9600
Protocol:
Template: 30-100 ng
Primer: each 1 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 10 ul
Buffer:
MgCl2: 1.5 mM
KCl: 100 mM
Tris-HCl: 10 mM
NH4Cl: 5 mM
pH: 8.6
The sequence for this STS was derived from a single sequencing
read. For additional information about the NHGRI chromosome 7
mapping project, see http://www.nhgri.nih.gov/DIR/STB/CHR7. Also
see Genomics 11:548-64 (1991) [MUID=92128937].
FEATURES
source Location/Qualifiers
1..245
/organism="Homo sapiens"
/map="7"
24..185
STS primer_bind 24..41
```

[illegible]

~~The sequence~~ for this STS was derived from a single sequencing read. For additional information about the NHGRI chromosome 7 mapping project, see <http://www.nhgri.nih.gov/DIR/CH7>. Also see Genomics 11:548-64 (1991) [MUID-92128937].

[illegible]

[illegible]

```

zr92e05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone 683168 3'.
DEFINITION
ACCESSION   AA214564
NID         g1813189
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.
REFERENCE   1 (bases 1 to 535)
AUTHORS     NCI-CGAP.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel.: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LLMU ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1129 Std Error: 0.00
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 275.
            Location/Qualifiers
FEATURES             source
                     1..535
                     /organism="Homo sapiens"
                     /note="Vector: pT73D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not 1; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human consillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TCTTACCAATCGAAGTGGGAGCGCGCTCATTTTTTTTTTTTTT-
                     3']. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT73 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
                     /clone="683168"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     complement(<1..>535)
                     /db_xref="GDB:5586488"
BASE COUNT      153 a 123 c 101 g 151 t 7 others
ORIGIN

Query Match      5.8%; Score 27; DB 14; Length 535;
Best Local Similarity 88.6%; Pred. No. 2.44e-02;
Matches 31; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 183 aaaaaaaagacagctattttattgtgaatatataa 217
      ||||| ||||| ||||| ||||| ||||| |||||
Qy 224 AAAAAAAAAAGAAAGATATATTATTGTAAGATAA 258

RESULT 13
ID HS1284157 standard; RNA; EST; 154 BP.
AC AA493647;
NI 9223488
DT 28-JUN-1997 (Rel. 52, Created)
DT 28-JUN-1997 (Rel. 52, Last updated, Version 1)
DE nr05c07.s1 NCI_CGAP_Thy1 Homo sapiens cDNA clone 943404.
KW EST.
OS Homo sapiens (human)
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
OC Homo.
RN [1]
RP 1-154
RT "National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

```


Search completed: Tue Feb 3 20:22:31 1998
Job time : 465 secs.

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.